


2019

Regulation of the long non-coding RNA FAM83H-AS1 by human papillomavirus in cervical cancer

Jamie Ann Barr Ph.D.

West Virginia University, jbarr3@mix.wvu.edu

Follow this and additional works at: <https://researchrepository.wvu.edu/etd>

 Part of the [Cancer Biology Commons](#), [Cell Biology Commons](#), [Molecular Biology Commons](#), [Oncology Commons](#), and the [Virus Diseases Commons](#)

Recommended Citation

Barr, Jamie Ann Ph.D., "Regulation of the long non-coding RNA FAM83H-AS1 by human papillomavirus in cervical cancer" (2019). *Graduate Theses, Dissertations, and Problem Reports*. 3926.
<https://researchrepository.wvu.edu/etd/3926>

This Dissertation is protected by copyright and/or related rights. It has been brought to you by the The Research Repository @ WVU with permission from the rights-holder(s). You are free to use this Dissertation in any way that is permitted by the copyright and related rights legislation that applies to your use. For other uses you must obtain permission from the rights-holder(s) directly, unless additional rights are indicated by a Creative Commons license in the record and/ or on the work itself. This Dissertation has been accepted for inclusion in WVU Graduate Theses, Dissertations, and Problem Reports collection by an authorized administrator of The Research Repository @ WVU. For more information, please contact researchrepository@mail.wvu.edu.

**Regulation of the long non-coding RNA
FAM83H-AS1 by human papillomavirus in cervical cancer**

Jamie A. Barr

Dissertation submitted
to the West Virginia University School of Medicine
at West Virginia University
in partial fulfillment for the degree of

Doctor of Philosophy
in
Cancer Cell Biology

J. Michael Ruppert, MD, PhD, Chair
Laura F. Gibson, PhD
John B. Barnett, PhD
Peter Stoilov, PhD
William P. Petros, PharmD

Ivan Martinez, PhD, Mentor

Cancer Cell Biology Program

Morgantown, WV
2019

Key Words: FAM83H-AS1, long non-coding RNA, lncRNA, human papillomavirus, HPV, cervical cancer

Copyright 2019 Jamie A. Barr

Abstract

Regulation of the long non-coding RNA FAM83H-AS1 by human papillomavirus in cervical cancer

Jamie A. Barr

Non-coding RNAs (ncRNAs), such as long non-coding RNAs (lncRNAs) and microRNAs (miRNAs), have been found to be involved in a variety of critical biological processes, and dysregulation of ncRNAs have been involved with several human diseases including cancer.

High-risk human papillomavirus (HPV) infection is one of the first events in the process of carcinogenesis in cervical and a subset of head and neck cancers. The expression of the viral oncoproteins E6 and E7 is essential in this process by inactivating the tumor suppressor proteins p53 and Rb, respectively, in addition to their interactions with other host proteins and regulation of ncRNAs. Our group identified novel regulation of host lncRNAs by HPV oncoprotein E6. More specifically, we discovered that a lncRNA known as FAM83H-AS1 is involved with proliferation, migration, and apoptosis in cervical cells, and high expression of this lncRNA correlates with poor overall cervical cancer patient survival. FAM83H-AS1 is a nuclear RNA, and mechanistically it is regulated through the E6-p300 pathway in a p53-independent manner. These findings provide knowledge of a specific lncRNA that could be studied further as a biomarker and/or therapeutic target not only in HPV-related cancers but also in other types of cancers where FAM83H-AS1 expression is dysregulated.

In parallel with these studies, our group identified a specific subgroup of miRNAs that are induced during quiescence and processed by a non-canonical biogenesis pathway by using primary human cells. miRNA expression is dysregulated when cells undergo a reversible state of growth arrest known as quiescence. These primary (pri-)miRNAs are modified with a 2,2,7-trimethylguanosine (TMG)-cap such that they are processed downstream in an Exportin-1 (XPO1)-dependent manner, independent of the canonical Exportin-5 (XPO5) protein used for exportation to the cytoplasm. The discovery of a new alternative miRNA pathway in quiescent primary human cells opens the door to future studies in other types of cells, such as stem cells and cancer stem cells, where the state of quiescence is important in their biological functions.

Dedication

This dissertation is dedicated to my parents, Ron and Regina Barr, who raised me to be hard-working and taught me persistence to strive to accomplish my goals while still enjoying the journey along the way. I am truly blessed to have such loving parents who have supported me through every aspect of my life. This dissertation work would not have been possible without them. I also want to thank my loving fiancé, Chris Juric, who provided constant optimism, laughter, and happiness when life was challenging, and motivation to continue working hard to complete this work. This dissertation is also dedicated to my Godparents, Larry Barr and Carol Pacofsky, who provided multiple care packages to brighten my day and reassurance that everything would work out.

I would also like to dedicate this dissertation to my supportive friends and family who were personally affected by cancer, especially my grandmothers, Irene Barr and Filomena Pacofsky, and grandfather, Larry Barr. Their experiences instigated my desire to enter into oncology research, and they kept me motivated when experiments failed. I am blessed to have the love and support of all of my friends and family, and I will forever be grateful for their contributions.

Acknowledgements

I would like to take this opportunity to thank my mentor, Dr. Ivan Martinez, for his guidance, constant support, and optimism throughout my graduate career. Being his first graduate student has provided us the opportunity to learn and grow together, therefore making us closer than the typical graduate student-principal investigator relationship. Looking back on how I progressed from a naïve to experienced scientist, I gained a multitude of knowledge from him, which made that transition and this dissertation possible.

I would also like to thank the Post-Doctoral Fellow, Karen Hayes, who was in the Martinez Lab for the bulk of my PhD years. She provided an unending amount of support both in and out of the laboratory, not only while she was a member of the lab, but even after she moved on to a new position. Her optimism, love for research, technical advice, and desire to help me succeed aided significantly in the completion of my dissertation. Additionally, my lab mates Tayvia Brownmiller and Abby Harold (e.g. the CRISPR whisperer) collaborated with me on multiple research projects and provided friendship in and out of the lab. We helped each other out along the way by sharing lab responsibilities, which allowed me to progress my work easier.

Lastly, I would like to acknowledge my committee members, Dr. J. Michael Ruppert, Dr. Laura F. Gibson, Dr. John B. Barnett, Peter Stoilov, and William P. Petros for their guidance and expertise throughout my PhD dissertation work.

Table of Contents

Abstract.....	ii
Dedication.....	iii
Acknowledgements.....	iv
Table of Contents.....	v
List of Main Figures.....	vii
 Chapter 1: Introduction and Literature Review.....	1
I. Non-coding RNAs (NcRNAs).....	1
<i>NcRNAs - Overview.....</i>	<i>1</i>
II. Small non-coding RNAs.....	4
<i>A. RNAi - Overview.....</i>	<i>4</i>
<i>B. miRNA Biogenesis - Canonical pathway.....</i>	<i>5</i>
<i>C. miRNA Biogenesis - Alternative pathways.....</i>	<i>7</i>
<i>D. miRNAs - Involvement in Cancer.....</i>	<i>11</i>
<i>E. miRNAs - Clinical Implications.....</i>	<i>14</i>
III. Long non-coding RNAs (lncRNAs).....	15
<i>A. LncRNA - Overview.....</i>	<i>15</i>
<i>B. LncRNA - Characterization of lncRNA.....</i>	<i>15</i>
<i>C. Circular ncRNAs - Overview.....</i>	<i>26</i>
<i>D. LncRNAs - Clinical implications.....</i>	<i>28</i>
IV. Human Papillomavirus (HPV).....	30
<i>A. HPV - Overview.....</i>	<i>30</i>
<i>B. HPV - Viral Oncoproteins.....</i>	<i>31</i>
V. LncRNAs in cervical cancer.....	33
<i>A. Cervical cancer - Overview.....</i>	<i>33</i>
<i>B. Host lncRNAs altered in cervical cancer.....</i>	<i>34</i>
VI. LncRNAs regulated by HPV.....	41
<i>A. HPV-regulated host lncRNAs.....</i>	<i>41</i>
<i>B. HPV regulation of lncRNAs in cancer - Clinical implications.....</i>	<i>43</i>
VII. Oncogenic lncRNA FAM83H-AS1 (also known as 'onco-lncRNA-3').....	43

Chapter 2: Long non-coding RNA FAM83H-AS1 is regulated by human papillomavirus 16 E6 independently of p53 in cervical cancer cells.....	63
Chapter 3: An Exportin-1 dependent microRNA biogenesis pathway during human cell quiescence	118
Chapter 4: Overall Discussion and Future Directions.....	186
Appendix: Immunoprecipitation of Tri-methylated Capped RNA.....	204
Curriculum Vitae.....	216

List of Main Figures:

Chapter 1: Introduction and Literature Review

Figure 1: Types of non-coding RNAs.....	1
Figure 2: Non-coding gene abundance with evolutionary organisms.....	2
Figure 3: Molecular mechanisms of lncRNAs.....	3
Figure 4: Canonical miRNA biogenesis pathway.....	6
Figure 5: Alternative miRNA biogenesis pathway in quiescent cells.....	10
Figure 6: Classification of lncRNAs by genomic location.....	17
Figure 7: lncRNA functions in the nucleus and cytoplasm.....	23
Figure 8: Circular RNA functions.....	27
Figure 9: HPV-mediated progression to cancer.....	30
Figure 10: HPV E6 oncoprotein regulation.....	32
Figure 11: HPV E7 oncoprotein regulation.....	33
Figure 12: lncRNA functions in cervical cancer.....	38
Figure 13: Heat map of FAM83H-AS1 expression alteration in multiple cancer types.....	47

Chapter 2: Long non-coding RNA FAM83H-AS1 is regulated by human papillomavirus 16 E6 independently of p53 in cervical cancer cells

Figure 1: Differential expression of host lncRNAs after expression of HPV-16 E6 in primary foreskin keratinocytes.....	82
Figure 2: Increased FAM83H-AS1 expression in primary cervical keratinocytes containing the HPV-16 genome as well as in HPV-16 positive cervical cancer and HNSCC lines.....	82
Figure 3: Regulation of FAM83H-AS1 expression by HPV-16 E6 in a p53-independent, p300-dependent manner.....	83
Figure 4: FAM83H-AS1 is localized in the nucleus in cervical pre-malignant and cancerous cell lines.....	84
Figure 5: FAM83H-AS1 knockdown altered cell proliferation, migration, and apoptosis in CaSki cells.....	84
Figure 6: FAM83H-AS1 expression is increased in human cervical cancer tissues and correlates with poor overall survival.....	85

Chapter 3: Exportin-1 dependent microRNA biogenesis pathway during human cell quiescence

Figure 1: Differential expression of mature miRNAs and their corresponding pri-miRNAs during quiescence.....	144
Figure 2: Regulation of Exportin-5 during quiescence by autophagy and miR-34a.....	145
Figure 3: Biogenesis of quiescence-induced miRNAs required for Exportin-1 but not Exportin-5.....	146
Figure 4: Quiescence-induced primary miRNAs are (TMG)-capped by TGS1.....	147
Figure 5: Detection of cytoplasmic pri-miR-34a is dependent on Exportin-1 expression.....	148
Figure 6: Detection of a small cytoplasmic isoform of Drosha during quiescence.....	149

Chapter 4: Overall Discussion and Future Directions

Figure 1: Heat map of FAM83H-AS1 expression alterations in multiple cancer types.....	189
Figure 2: HPV-16 was the only high-risk strain analyzed that demonstrated E6 oncoprotein binding to p300.....	192
Figure 3: Schematic of SAM complex.....	195
Figure 4: CRISPR Cas9 mutant technology to over-express FAM83H-AS1 in C-33A cell line.....	196

Appendix: Immunoprecipitation of Tri-methylated Capped RNA

Figure 1: Main steps diagram of this protocol.....	212
Figure 2: RNA immunoprecipitation of (TMG)-capped primary miRNAs (pri-miRNAs) in quiescent human foreskin fibroblasts (HFFs).....	213

Chapter 1

Introduction and Literature Review

I. Non-coding RNAs (NcRNAs)

NcRNAs - Overview

In 1961, the concept of messenger RNAs (mRNAs) was determined as the intermediate messenger between DNA and protein synthesis(1). This led to the central dogma of biology that DNA containing the genetic information is transcribed to mRNA, which in turn is translated to a protein; the protein is considered the critical functional product. Non-coding RNAs contribute to the translational process, including ribosomal RNA (rRNA) and transfer RNA (tRNA) processing. rRNA associates with proteins to form ribosomes, which catalyze the assembly of amino acids into protein chains. Each amino acid is covalently bound to a specific tRNA, which guides it to the polypeptide chain and helps decipher the next desired amino acid to add to the chain(2).

In 1982, the first "enzymatically active" RNA was identified (ribozyme) with a specific biological

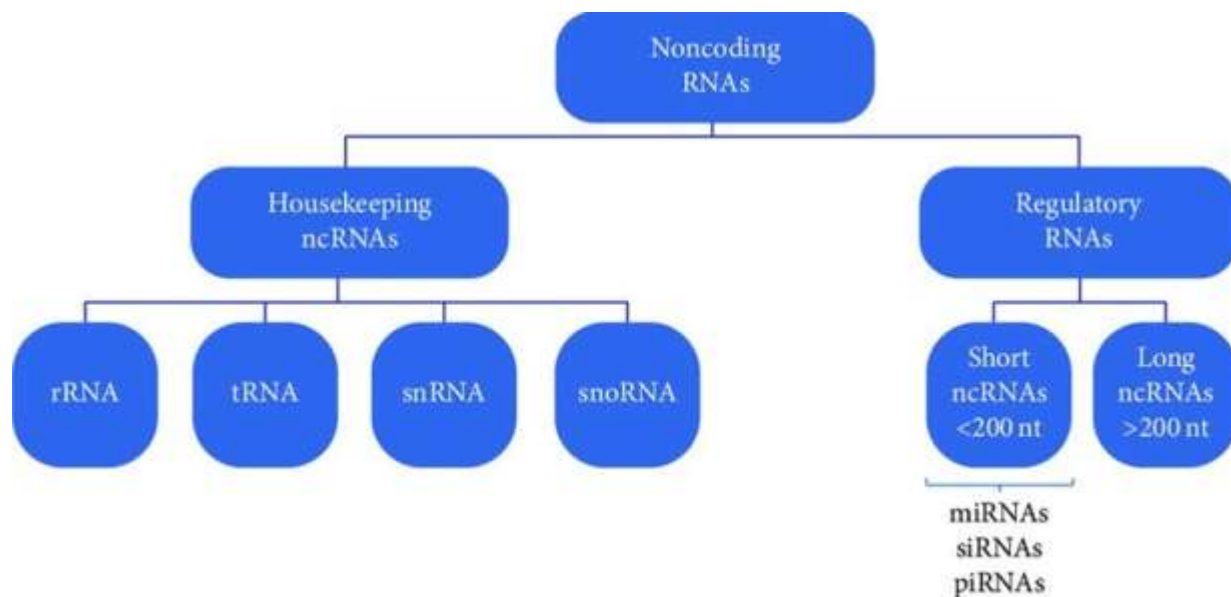


Figure 1: Types of non-coding RNAs.

Modified from: Losko et al., *Mediators of Inflammation*, 2016.

function(3). These RNA enzymes catalyze their own intronic excision during RNA splicing (self-splicing

introns)(4). These early findings opened the door to research additional RNAs exhibiting non-coding functions and, to date, there are various other functional housekeeping and regulatory ncRNAs summarized in Figure 1.

Advancement of sequencing technologies, the Encyclopedia of DNA Elements (ENCODE) work(5), and completion of the Human Genome Project revealed that ~80% of the human genome is transcribed and the majority of transcribed genes do not code for proteins (<3% of the human genome codes for proteins)(6). It was initially thought that the majority of the non-coding genomic regions were "transcriptional noise", but it is now appreciated that many of these have more biologically relevant functions.

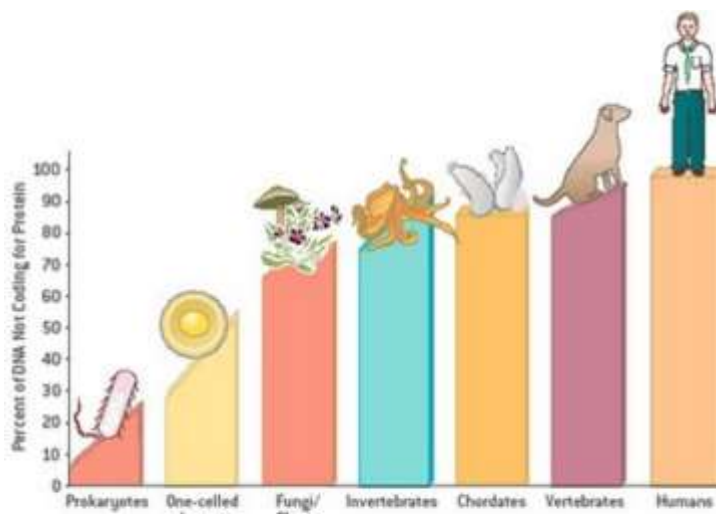


Figure 2: Non-coding gene abundance with evolutionary organisms.

Adapted from: Mattick JS, *Scientific American*, 2004.

These various ncRNAs once considered to be "junk" have been identified to rebel against the status quo and have other biological functions besides coding for proteins. NcRNAs typically lack an open reading frame (ORF) such that they do not code for proteins. Some ncRNAs contain very small ORFs that produce small peptides(7, 8), but this does not prove they are functional. More evolutionary complex organisms have a greater abundance of ncRNAs than less complex organisms

(Figure 2) leading towards the theory that ncRNAs contribute to the development and control of more complex cellular processes. Most ncRNAs form RNA-protein complexes called ribonucleoproteins (RNPs), however, they can have a multitude of alternative interactions. NcRNAs can also act as competing endogenous RNAs (ceRNAs) such that the ncRNA can mimic an RNA binding partner and inhibit the function of the alternative binding partner interaction. NcRNAs are typically classified based on their size, into small and long non-coding RNAs (lncRNAs). Common small ncRNAs include small nuclear RNAs (snRNAs), small nucleolar RNAs (snoRNAs), microRNAs (miRNAs), small interfering RNAs (siRNAs), and PIWI-interacting RNAs (piRNAs) (Figure 1).

In mRNA processing, pre-mRNA is the initial product transcribed, which contains both exons and introns. In 1993, a Nobel Prize in Physiology and Medicine was earned for identifying that introns can be

spliced out by spliceosomes (Philip Sharp and Richard Roberts). The spliceosomes are composed of several small nuclear (sn) RNP-specific proteins interacting with non-coding snRNAs. The snRNA can form an RNA-RNA pair with pre-mRNA in the required alignment for downstream splicing steps(9).

RNA interference (RNAi) is the process of RNAs directing enzymes to mRNAs to induce post-transcriptional gene silencing and therefore regulate vital cellular processes. The first RNAi type of phenomenon was reported in 1990 in plants(10). In 1998, RNAi was identified in *C. Elegans* and they further identified that it was double-stranded RNAs triggering gene silencing(11). In 2006, Andrew Z. Fire and Craig C. Mello were awarded the Nobel prize for discovering RNAi. Two central types of ubiquitously expressed RNAi are miRNAs and siRNAs, while piRNAs are another form of RNAi that are expressed mainly in germline cells(12). Pseudogenes resemble functional genes, but have coding-sequence deficiencies. RNAi genes, such as siRNA(13), can derive from pseudogenes. Additionally, pseudogene ncRNAs can act as RNAi, such as miRNA sponges, to regulate gene expression(14).

Whereas RNAi mechanisms of action in gene regulation are well understood, lncRNAs are not yet well characterized. LncRNAs are defined as RNAs greater than 200 nucleotides in length that are not translated to proteins. In 2002, lncRNAs were officially identified as a significant new class of non-protein coding RNA molecules after the sequencing of a full-length cDNA library in mouse(15). Since then, they have been found to function as signal, guide, decoy, and scaffold RNAs (Figure 3).

All of these types of ncRNAs have been found to play roles in a variety of pathological conditions, including cancer. Specific ncRNAs expression has been correlated with survival and therefore are being studied as biomarkers for disease diagnosis and prognosis. Additionally,

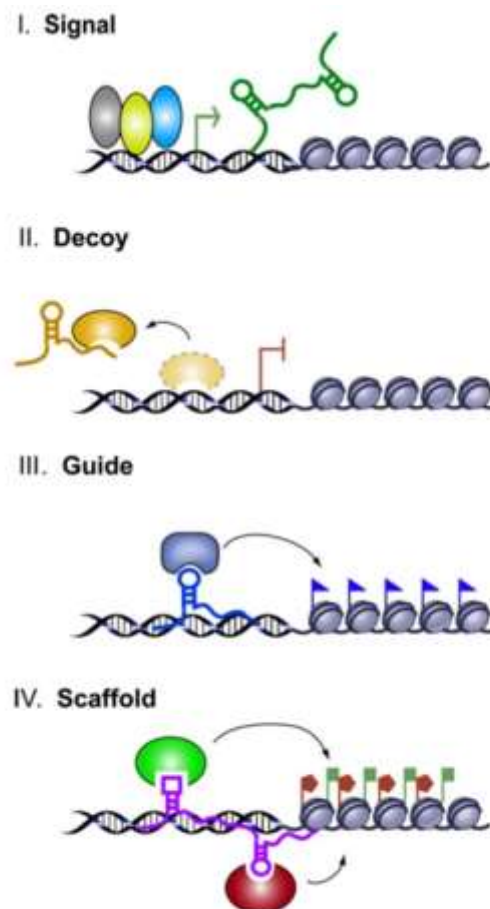


Figure 3: Molecular mechanisms of lncRNAs.
Adapted from: Wang and Chang, *Molecular Cell*, 2011.

specific ncRNAs have been identified to be functionally related to hallmarks of cancer and are being pursued as therapeutic targets.

II. Small non-coding RNAs

A. RNAi - Overview

In 1993, the first microRNA, *lin-4*, was identified in *C. Elegans*, by the Ambros and Ruvkun labs(16). However, miRNAs were not respected as functionally critical until 2002 when Carlos Croce at the Kimmel Cancer Institute (now at The Ohio State University) identified, for the first time, two miRNAs involved in human cancer(17). Around that time, hundreds of other miRNAs were reported and discovered to be conserved amongst species(18). These small ncRNAs were initially referred to as "small temporal RNAs (stRNAs)" prior to 3 publications in *Science* in 2001 that transitioned the field towards the term "microRNAs"(19-21).

MicroRNAs are short, single stranded ncRNAs that are ~22 nucleotides long. They function by binding to the miRNA response elements (MREs) in the 3' untranslated region (3'-UTR) of mRNAs and targeting them for degradation and/or preventing their translation. MiRNAs contain an eight-base seed region that is well conserved and binds with perfect complementarity to the mRNA target. miRNAs with similar seed sequences with likely target a similar set of genes, however miRNAs with different seed sequences can still target and regulate similar gene sets.

miRNAs have been found to be well conserved in many species, including humans, and are master regulators of almost every biological process. In humans, it is predicted that miRNAs can regulate more than 90% of genes, and a single miRNA can target multiple genes(22). They are involved with immunity, development, and metabolism, and dysregulation of these miRNAs is implicated in most human diseases, including cancer, cardiovascular diseases, and neurological diseases. Currently, there are 2,654 annotated human miRNAs(23).

In 1999, a similar ncRNA known as small interfering RNA (siRNA) was identified in plants(24). A long double-stranded RNA is cleaved by the endonuclease Dicer to form a small interfering RNA (siRNA) that is ~20-25 nucleotides in length. SiRNA can enter into a cell, get incorporated into other proteins to form the RNA-Induced Silencing Complex (RISC). At this point, the siRNA is unwound and separated to form a single stranded siRNA. The single-stranded siRNA that remains part of the RISC

complex then binds to complementary target mRNA to induce mRNA cleavage, and leads to mRNA degradation.

miRNAs and siRNAs show many similarities, but there are some distinctive traits between the two RNAs. siRNAs are derived from longer regions of double-stranded RNA. Both siRNA and miRNA are similar in length and mediate downregulation of their specific mRNA target. miRNAs are regulators of endogenous genes, while siRNAs regulate in response to foreign entities (e.g. viruses). It is suggested that siRNA destabilizes mRNA, while miRNA primarily inhibits mRNA translation without affecting the mRNA level. miRNA can sequester and degrade mRNA by P-bodies, which are sites for RNA decay(25). The main difference between siRNA and miRNA is that siRNA has perfect complementarity to its target, while most miRNA only has complementarity in the seed region with variable complementarity of the target to the rest of the miRNA. This feature allows miRNAs to have greater plasticity to target several mRNAs, while siRNAs only can target mRNA that have perfect complementarity. Additionally, miRNAs and siRNAs can be synthetically developed and utilized as a research tool as well as clinical therapeutic applications.

More recently, another form of RNAi known as piRNAs were identified to prevent genome instability in germ line cells. piRNAs interact with piwi proteins to form RNA-protein complexes that are involved in post-transcriptional silencing of transposons and other cellular genes. piRNAs are transcribed from single-stranded precursor transcripts that are processed in a Dicer-independent mechanism and are slightly longer (~24-31 nucleotides) than miRNAs and siRNAs (12).

B. MiRNA Biogenesis - Canonical Pathway

The canonical biogenesis of miRNAs has been elucidated as a stepwise process in normal and aberrant cells (Figure 4). The initial step of miRNA biogenesis is transcription of a long primary (pri)-miRNA, typically by RNA Polymerase II (Pol II)(26). Some miRNAs are transcribed by RNA Polymerase I (Pol I). The majority of pri-miRNA are transcribed from intergenic regions or antisense to host genes while others are transcribed from intronic regions; thus, miRNA may be formed independently or as part of host genes. The size of a pri-miRNA ranges from several hundred nucleotides to several kilobases, and contains a stable 7-methylguanosine (m7G) cap at its 5' end and polyadenylated cap at its 3' end. These pri-miRNAs contain one or more bulged structures. The canonical biogenesis pathway then consists of two cleavage events, one nuclear and one cytoplasmic, to produce a mature miRNA.

In the nucleus, the microprocessor consisting of the RNase III nuclease Drosha(27) and double stranded RNA binding protein DiGeorge Syndrome Critical Region 8 (DGCR8)(28) recognizes the bulged hairpin structure of the pri-miRNA and cleaves the pri-miRNA to precursor (pre-) miRNA. This leaves the pre-miRNA with a ~70 nucleotide stem-loop structure. More specifically, Drosha cleavage generates 3' protruding ends that are later recognized by Exportin-5 (XPO5) and Dicer(27, 29, 30).

Notably, RNase III protein homologs conservation and interactions can vary between species. Initially Pasha was identified as Drosha's partner required for miRNA maturation and miRNA-mediated gene regulation in *Drosophila* and *Caenorhabditis elegans* (*C. Elegans*)(31) and it was later identified as DGCR8 in other organisms(28).

Pre-miRNA intermediates are bound by XPO5 and exported from the nucleus to the cytoplasm(30, 32, 33) in a Ran-GTP dependent manner. Subsequently, pre-miRNAs are cleaved by the cytoplasmic RNase III protein Dicer1 and Dicer-transactivating response RNA-binding protein (TRBP) complex into ~22 nucleotide miRNA duplexes(34-38). This miR/miR* (miR-3p/miR-5p) duplex is then unwound by a helicase. One strand of the cleaved duplex, sometimes referred to as the "passenger strand" (miRNA*), can be non-functional and degraded by a nuclease (39). The other "guide strand" preferentially remains a functional mature miRNA(40, 41). The miRNA* terminology is now used less, however, because some

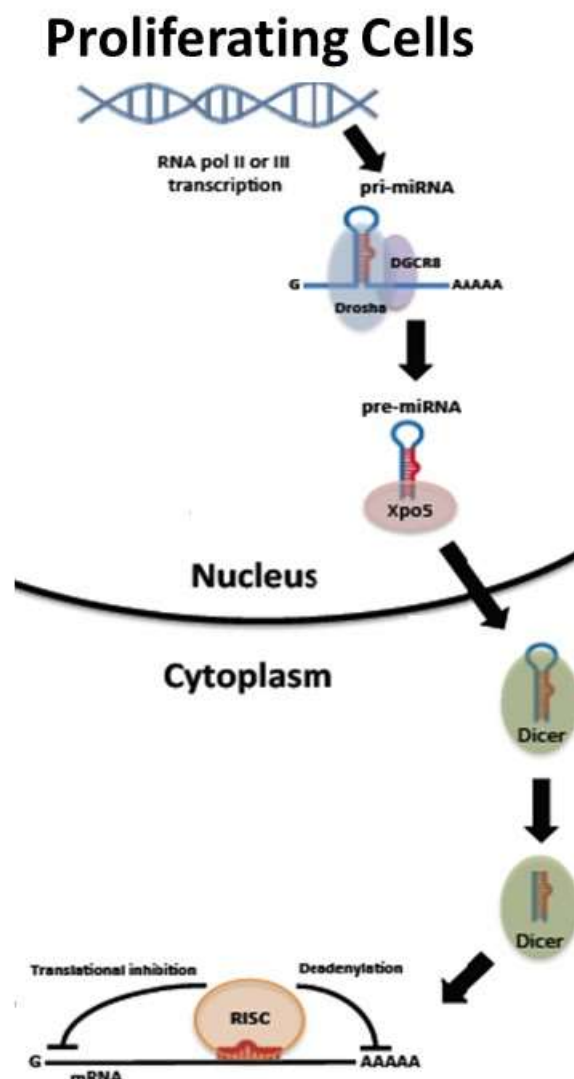


Figure 4: Canonical miRNA biogenesis pathway.
Adapted from: Martinez et al., *PNAS*, 2017.

miRNA* sequences were found to be functional. Currently, the two miRNA strands are named miR-3p and miR-5p based on the direction of the strand it derived from, and both strands have the potential to be functional(42). The single-stranded mature miRNA guide strand is loaded into the Argonaute-2 (Ago2) containing RNA-induced silencing complex (RISC). Each RISC can only contain one of the two strands. The stability of the base pairs at the 5' end of each predetermines which strand is loaded into the RISC (41); the strand with the less stable 5' end usually survives. The mature miRNA contained in the RISC will then base-pair incompletely with the target mRNA. Depending on the degree of complementarity between the miRNA and mRNA target will determine the fate of the mRNA. Most miRNAs exhibit imperfect complementarity with target mRNA and will inhibit the mRNA's translation(43) by inhibiting the ribosome progression. Perfect complementarity will induce degradation of the mRNA(44). For example, miR-196 is perfectly complementary to its mRNA target HOXB8 (except for a wobble base pairing) and the miRNA induces degradation of the mRNA(44). miRNA binding destabilizes the mRNA by speeding up the deadenylation of the mRNA and contributing to mRNA degradation(45).

C. miRNA Biogenesis- Alternative Pathways

Even though the canonical pathway is well studied, deep sequencing of systems with deletions of miRNA biogenesis pathway components have identified non-canonical biogenesis pathways in different organisms and cells that can still produce functional miRNAs. These alternative pathways can differ in the number of cleavage events and enzymes responsible for miRNA processing. It appears that the biogenesis pathway that a miRNA takes to be processed is determined by its genomic site of origin, sequence, and thermodynamic stability.

In the same human cell line, Drosha, XPO5, and Dicer were deleted to analyze for effects on miRNA biogenesis. Drosha-depleted cells exhibited completely abolished canonical miRNA biogenesis. From this analysis, they detected three miRNAs processed independent of Drosha (miR-7706, miR-3615, and miR-1254), while canonical miRNA production was completely abolished in cells with Drosha depletion. Dicer depletion led to reduced levels canonical miRNAs, but many were still detected. In cells with Dicer knockdown, pre-miRNAs were loaded directly onto Argonaute protein and trimmed at the 3' end to produce miRNAs from the 5' strand (5p miRNAs). Unexpectedly, XPO5 depletion only partially affected miRNAs expression(46), suggesting there are alternative miRNA exportation pathways. This study contributes to the idea that certain miRNAs can be processed independent of the main proteins involved with canonical biogenesis.

Mirtrons

In 2007, the Bartel and Lai laboratories described a group of RNAs in *Drosophila melanogaster*(47, 48) and *Caenorhabditis elegans* (*C. Elegans*)(48) termed "mirtrons" (pre-miRNA/intron) that follow a non-canonical miRNA biogenesis pathway as they are Drosha- and DGCR8-independent. Similar to canonical miRNAs, mirtrons are encoded within short stem-loop structures. However, while miRNAs utilize the microprocessor complex for processing, mirtrons are instead generated from the excised intron of protein-coding genes that are released during mRNA splicing. More specifically, a small intron lariat is linearized by the lariat debranching enzyme to produce a pre-miRNA-like short hairpin. Thus, mirtrons are processed independently of the microprocessor complex, but follow the canonical miRNA pathway during nuclear export and thereafter(47, 48). Subsequently, mirtrons were found to be well-conserved and expressed in diverse vertebrates(49-52).

With conventional mirtrons, both of their ends are defined by splicing; however, some mirtrons contain 5'-(52) or 3'-(52, 53) single-stranded RNA tails that need to be trimmed by nucleases before Dicer processing. For example, certain 3'-tailed mirtrons (e.g. hsa-miR-4745 and hsa-miR-6869) are trimmed by the RNA exosome(52, 53).

Simtrons

While mirtrons are splicing-dependent, a subset termed "simtrons" (splicing-independent mirtron-like miRNAs) do not require DGCR8, Dicer, Exportin-5, or Argonaute 2 (AGO2). They can however interact with Argonaute proteins. They are exported from the nucleus in an unknown mechanism. They are bound by, and dependent on, Drosha. Simtrons function by silencing target transcripts and are found in the RISC complex. Some examples are miR-1225 and miR-1228(54).

snoRNA-miRNAs

In 2008, Ender C. et al. discovered small RNAs derived from snoRNA ACA45 that are processed independent of the microprocessor Drosha/DGCR8 complex and function like miRNAs. The processing of this RNA is still dependent on Dicer and interacted with Ago proteins to silence mRNA (CDC2L6) expression post-transcriptionally in a miRNA-like manner(55).

Since then, other snoRNAs with dual roles as miRNAs (e.g. miR-140, miR-605) have been identified. Based on sequence and structure, there are two classes of snoRNAs (box H/ACA and box C/D), both of which produce snoRNA-derived RNAs(55-59).

miRNAs from tRNAs

In 2008, Babiarz et al. identified a pre-tRNA (miR-1983) that was Dicer-dependent and DGCR8-independent(50). Later studies demonstrated other tRNA-derived small RNAs (tsRNAs) involved in gene silencing that were Dicer-dependent (type I tsRNAs) (60, 61), as well as some tsRNAs processed by RNase Z cleavage downstream from a tRNA gene in a Dicer-independent manner (type II tsRNAs)(61). These tsRNAs localize to the cytoplasm and associate with Argonaute proteins(61).

Virus-derived miRNAs

Other tRNA-derived small RNAs that were miRNA-like were identified to be generated from both host and viral genomes(62-64). Murine gamma-herpesvirus 68 (MHV68) miRNAs are co-transcribed fused to a tRNA from RNA polymerase III promoters(63). The resulting pri-miRNA is then processed by tRNase Z, independent of Drosha, to form a pre-miRNA that then exhibits canonical processing to yield mature viral miRNAs (e.g. MHV68 miR-M1-7)(64).

In certain instances, potentially to combat viral infection (e.g. Sinbis RNA virus), pri-miRNAs can be incorporated into a cytoplasmic virus. The pri-miRNA is downstream processed in a Dicer-dependent, DGCR8-independent manner to produce a viral mirtron-like product (virtron)(65). Following viral infection, Drosha is relocalized to the cytoplasm to aid in this processing of virus-derived miRNAs(66).

Expression of viral miRNAs from herpesvirus saimiri (HVS) relies on elements required for the synthesis and processing of Sm class U RNAs known as HVS U RNAs (HSURs). More specifically, these viral miRNAs are processed independent of the Microprocessor complex. Instead the Integrator complex (mediates processing of snRNA) cleaves to generate the 3' end of the HSUR and the pre-miRNA hairpin. To generate these mature viral miRNAs, the pre-miRNA is then processed downstream by the canonical pathway involving XPO5 and Dicer(67).

Recently, viral miRNAs produced from certain low risk strains of HPV (17, 37, and 41) were identified. Further studies found that these viral miRNAs are processed by canonical miRNA biogenesis and can regulate viral transcripts (e.g. E1) and consequently control their own viral life cycle(68).

Dicer-independent miRNA biogenesis

miR-451 is a conserved vertebrate miRNA that is cleaved by Drosha/DGCR8 to generate a pre-miRNA that is too short to be a substrate for Dicer. Instead, the pre-miRNA is loaded directly into Argonaute proteins. If the pre-miRNA is loaded into an Ago2 protein, it is cleaved(69-72).

XPO1-dependent miRNA biogenesis

Previously, capped RNAs shorter than 200 nucleotides were found to be bound by CBC (cap-binding complex)-PHAX and exported via XPO1 rather than XPO5.(73) In a similar manner, certain pre-miRNAs (e.g. pre-miR-320 and pre-miR-688) were found to be processed differently with regards to pre-miRNA synthesis, nuclear-cytoplasmic transport, and guide strand selection in mice. They have a 7-methylguanosine (m⁷G)-cap employed by eIF4E that is recognized by XPO1 via the adapter molecule PHAX. These certain capped pre-miRNAs were processed independently of the Microprocessor. Once exported in an XPO5-independent manner, they are cleaved by Dicer and only the 3p-miRNA is efficiently loaded onto Argonaute to form mature, functional miRNA(74).

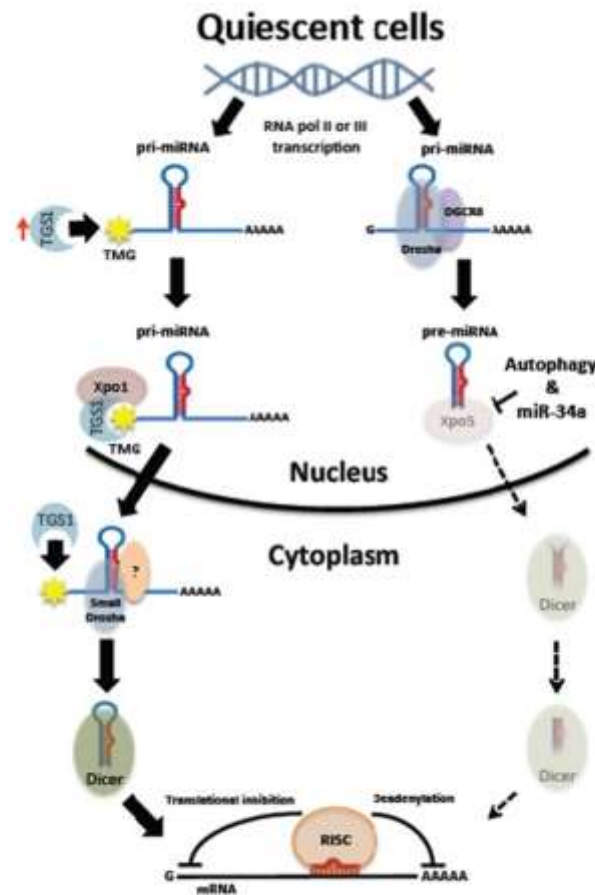


Figure 5: Alternative miRNA biogenesis pathway in quiescent cells.
Adapted from: Martinez et al., *PNAS*, 2017.

The Martinez lab recently discovered another XPO1-independent, non-canonical pathway of miRNA biogenesis in normal, human foreskin fibroblast (HFF) quiescent cells (Figure 5). XPO5 expression is lost in quiescent cells by autophagy and the regulation of miR-34a. A group of miRNAs are induced during quiescence in HFFs and their biogenesis is XPO1-dependent (XPO5-independent). These pri-miRNAs (e.g. pri-miR-34a and pri-miR-3188) are modified with a 2,2,7-trimethylguanosine (TMG)-cap by trimethylguanosine synthase 1 (TGS1), which allows them to bind to and be exported by XPO1. Additionally, pri-miR-34a was present in the cytoplasm along with a small isoform of Drosha, implying further potential alternative processing of this miRNA in quiescent cells.

In summary, the importance of miRNAs is essential in so many processes that evolution gives the plasticity to process these miRNAs in non-canonical pathways as needed.

D. miRNAs - Involvement in Cancer

MiRNAs involvement in cancer has been comprehensively reviewed(75-78). Global differential expression analyses of miRNAs in human cancer cells and tissues have revealed dysregulation of many miRNAs in human cancer and that most miRNAs are repressed in cancer tissues compared to normal(79, 80). Global depletion of miRNAs via impaired miRNA processing, led to cell transformation and tumorigenesis, which implies miRNA alteration contributes to cancer development(81). Functionally, miRNAs can be both tumor suppressive or oncogenic depending on the context, and are involved in many aspects of cancer biology, including proliferation and apoptosis. In many cases, the trend of miRNA expression correlates with many clinical variables (e.g. tumor stage and overall survival) such that miRNA profiling of a patient can lead to improved diagnosis and prognosis in certain cancers(77). Additionally, some have proposed that miRNA profiling is even better at predicting clinical variables than mRNA profiling in certain cases explored thus far(79).

Mechanisms of miRNA alterations in cancer

There are multiple mechanisms of miRNA deregulation in cancers, including genomic abnormality, epigenetic factors, transcriptional regulation, and miRNA biogenesis regulation as described below. Notably, stoichiometry analysis of miRNA to target gene abundance can allow for the speculation of how likely the miRNA is to regulate the target and contribute to a biological effect. Higher levels of mRNA expression results in varied downregulation of targets by miRNA(82).

The first example of a miRNA involved in cancer was miR-15a and -16-1, which are clustered at a region of the genome that is frequently deleted in B cell chronic lymphocytic leukemias (B-CLL). Thus, in the majority of CLL cases, both genes are deleted or down-regulated(17). Other groups have additionally showed deregulation of miRNAs in cancer due to chromosomal abnormalities (e.g. fragile sites)(83).

miRNA expression can be regulated by epigenetic factors, as many CpG islands are located near miRNAs(84). CpG hypermethylation repressed miRNAs (e.g. miR-9-1 and miR-124a) in cancer tissues compared to normal, and this miRNA hypermethylation affected known tumor suppressors and oncogenic factors (84, 85). Alternatively, miRNAs (e.g. miR-29) can counteract CpG methylation by directly targeting DNA methyltransferases and therefore de-repress downstream targets regulated by promoter methylation in cancer cells(86).

Well-known transcription factors involved in cancer can up-regulate the expression of certain miRNAs to contribute to carcinogenesis. For example, Myc(87) and p53(88) can alter expression of the miR-17-92 cluster and miR-34, respectively, in cancer cells.

Alterations in expression of mature miRNAs can occur due to decreased miRNA biogenesis processing efficiency or miRNA stability. In cancer, key regulators of the miRNA biogenesis pathway (Drosha or Dicer) have altered expression, which affected steady-state miRNA levels and led to tumor development(81).

In some cases, miRNA expression levels are not altered, but rather there is sequence variation in the miRNA binding sites in the target gene that affect miRNA-mRNA interactions and contribute to cancer(89).

Oncogenic or tumor suppressive miRNAs

Certain miRNAs are up-regulated in a cancer lineage-specific manner, while others are altered in many different types of cancers and thus critical tumor suppressors or oncogenes. Some of the most well-known miRNAs that are altered in many cancers from different lineages are miR-17-92 cluster, miR-21, miR-155, let-7, and miR34s(77). Since miRNAs target and regulate expression of mRNAs, a miRNA is determined to be oncogenic or tumor suppressive depending on its target gene.

The oncogenic miR-17-92 cluster (composed of miR-17, -18a, -19a, -20a, -19b-1, and -92a-1) is up-regulated in many human cancers(76). It is up-regulated by MYCN and binds the 3'UTR of the cyclin-dependent kinase inhibitor CDKN1A (p21) and pro-apoptotic regulator BCL2L11 (Bim) mRNAs in neuroblastoma(90).

miR-21 regulates multiple target genes, including p53 and NF-kB, involved in hallmarks of cancer(91). miR-21 expression is up-regulated in many different types of cancers(76) and is considered to be a ubiquitous oncogene as it is expressed in all epithelial-cell derived solid tumors and many other cells(91). In the IL-6 signaling pathway, Stat3 transcriptionally activates miR-21 such that downstream it regulates invasion and metastasis(92).

miR-155 has been identified as a bona fide oncogene in breast cancer because it binds key cancer targets (e.g. FOXO3A and RhoA) for degradation or inhibition of translation to accelerate tumor development(93). miR-155 has also been found to drive therapy resistance in various tumor types and is being studied as a target for miRNA-based therapeutics(94).

Let-7 is a tumor suppressor that targets well-known oncogenes RAS, c-Myc(81), and HMGA2(76). The expression of let-7 has been reported to be down-regulated in many human cancers, including lung and colon(95).

miR-34a acts as a tumor suppressor by regulating of at least 700 transcripts involved in cellular processes such as proliferation (cyclin-dependent kinases), apoptosis (BCL2), and motility (SNAIL). miR34a is transcriptionally regulated by the tumor suppressor p53 in normal conditions, however miR-34a expression is decreased in many cancers likely due to alterations in p53(96).

Virus-regulation of miRNAs

The involvement of miRNAs in the pathogenesis of oncogenic viruses was comprehensively reviewed by Lin et al. in 2011(97). Certain oncogenic viruses can alter host miRNAs(98, 99) while others can encode their own viral miRNAs(100, 101) to control their viral life cycle, but consequently this can contribute to carcinogenesis(97).

The high-risk human papillomavirus (HPV) E6 oncoprotein inhibits expression of the tumor suppressor miR-34a(99). Additionally, Martinez et al. showed that high-risk HPV E6 inhibits miR-218, which up-regulates LAMB3 to contribute to HPV's tumor promoting function(98).

Kaposi sarcoma-associated herpesvirus (KSHV)(100) and Epstein-Barr virus (EBV)(101), can produce viral miRNAs to regulate carcinogenesis. For example, a cluster of KSHV miRNA decreases an anti-angiogenic, tumor suppressor thrombospondin 1 (THBS1)(100, 102).

E. miRNA - Clinical Implications

Regulating miRNAs for application in a clinical setting has evolved by using antagomir and miRNA mimics. To counteract the up-regulation of oncogenic miR-17p-92 cluster in therapy-resistant neuroblastoma, *in vitro* and *in vivo* antagomir-17-5p treatment was used to abolish the growth of MYCN-amplified cancer cells by up-regulating p21 and BIM to blockade the cell cycle and activate apoptosis(90). A Phase 1 clinical trial was completed that tested intravenous injection of TargomiRs (targeted minicells containing a miRNA mimic) for the first time in humans. It was tested as a second- or third-line treatment for patients with recurrent malignant pleural mesothelioma (MPM) and non-small cell lung cancer (NSCLC). The TagomiR consisted of a synthetic miR-16 mimic, non-living bacterial minicells (nanoparticles), and an anti-EGFR specific antibody. At this time, no results have been posted for this specific study (NCT02369198).

Additionally, there is the potential of analyzing for expression of miRNA(s) for use as diagnostic and/or prognostic marker. There are various clinical trials underway to analyze circulating miRNAs in blood and plasma as biomarkers of different types of cancers to diagnose cancer and stratify patients by stage (to determine treatment). Certain Phase I clinical trials (e.g. NCT02509052). are combining their primary objectives with secondary and tertiary objectives to preliminarily analyze miRNA signatures associated with responses of cancer cells to therapeutic treatments. The Spanish Breast Cancer Research Group is sponsoring a clinical trial to analyze whole blood of metastatic breast cancer patients for a miRNA signature to predict response to bevacizumab therapy (NCT01598285). Analysis of ncRNA expression in exosomes is currently an actively pursued direction of clinical research. One clinical study is currently recruiting patients to analyze ncRNA expression via genetic sequencing of miRNA/lncRNA in the exosome of epithelia ovarian cancer patients. (NCT03738319). There are clinical trials researching miR-155 as a diagnostic marker in non-muscle invasive bladder cancer (NCT03591367). Notably, many clinical trials are looking for a miRNA signature of multiple miRNAs rather than relying on a single miRNA biomarker.

One specific colon cancer study is currently analyzing for expression of 6 miRNAs (miR-21, -20a-5p, -103a-3p, -106b-5p, -143-5p, -215) in clinical surgical specimens to identify if a patient should additionally receive adjuvant chemotherapy (NCT02466113). A Phase 4 clinical trial detecting fifteen circulating miRNAs (identified from preclinical studies) as biomarkers of hormone sensitivity in breast cancer patients before and after first-line therapy was completed. In parallel, they analyzed a larger scale analysis of circulating miRNAs in these patients (NCT01612871). Prospective studies are being conducted, such as profiling plasma miRNA expression in heavy smokers to assess its efficacy as a first-line screening for lung cancer screening of high-risk population (NCT02247453). There is a completed clinical trial analyzing for expression of miRNA biogenesis proteins (Dicer and Drosha) in epithelial skin cancer (NCT00849914). A clinical trial sponsored by MiraKind is validating the role of miRNA binding site mutations in cancer risk, prevention, and treatment (NCT02253251). Unfortunately, at this time, no results have been posted on clinicaltrials.gov for any of the studies mentioned above.

As miRNA mechanisms of action to function in gene regulation are well understood, lncRNAs are not yet well characterized.

III. Long non-coding RNAs (LncRNAs)

A. LncRNA - Overview

In 2009, Dr. John Rinn and Dr. Eric Lander's labs at the Broad Institute identified over a thousand lncRNAs that are highly conserved in mammals(103). This contributed to the great increase in research publications on lncRNAs in the past ~10 years(104). To date, there are currently over 127,000 annotated human lncRNA transcripts in the LNCipedia public database (lncipedia.org), however, few of these have been fully characterized. Admittedly, lncRNA transcription does not necessarily mean that it has a biological function, however many lncRNAs have been characterized as functionally important and it is critical to continue to dissect functions. LncRNAs make up the majority of the non-coding portion of the genome (<https://www.encodegenes.org/human/stats.html>), so it is critical to further research and characterize this class of RNAs.

B. LncRNA - Characterization of lncRNA

LncRNAs were initially identified as mRNA-like transcripts that do not code for proteins; however, further characterization of lncRNAs has distinguished them as their own class. It is still challenging to distinguish between coding and lncRNA transcripts.

Since lncRNAs as a class are not well characterized yet, there are only trends, not standard "rules", to classify lncRNA based on biogenesis, processing, localization, functions, etc. As expected, there are also lncRNAs that have been identified as exceptions to the trends of the majority of lncRNAs. A review in 2016 summarized certain research techniques that could be used to identify and characterize a newly identified lncRNA(105). Once a specific lncRNA is identified, researchers can determine: biogenesis/processing, genomic location, non-protein coding potential; conservation; expression level, size, and primary sequence; secondary structure; localization; and biological function.

LncRNA - Biogenesis and Processing

In terms of biogenesis, lncRNAs and mRNAs share similar characteristics. LncRNAs are typically transcribed by RNA Polymerase II (Pol II) from genomic loci with similar chromatin states to mRNAs(103). In many cases, lncRNAs differ from mRNAs solely in that lncRNAs are lacking a translated ORF. Similar to mRNAs, lncRNAs are modified with a 5' cap and 3' polyadenylated tail(103) and undergo pre-RNA splicing. Some lncRNAs undergo non-canonical processing at their 5' and 3' ends(106-108).

The promoter regions of the majority of lncRNAs display similar epigenetic marks [e.g. increased trimethylation of lysine 4 of histone 3 (H3K4me3)] to mRNAs. LncRNAs have higher enrichment of acetylation of lysine of histone 3 (H3K27ac) and are more strongly repressed by certain chromatin remodeling complexes (e.g. Swr1, Isw2) than mRNAs(106). Some eukaryotic promoters are bi-directional such that a sense protein-coding gene could be transcribed while an antisense is transcribed in the opposite direction(109). Polyadenylation signals differ between sense (protein coding) and antisense strands (lncRNA) produced from divergent promoters; polyadenylation usually occurs during splicing of mRNAs, while it typically occurs during cleavage in lncRNAs.

Similar to mature tRNAs, some lncRNAs (e.g. MALAT1) use an alternative processing method consisting of cleavage by the RNase P ribonucleoprotein complex. These specific lncRNAs contain a tRNA-like structure at their 3' ends that RNase P cleaves to produce mature lncRNA transcripts. After RNase P cleavage, the mature lncRNA has an RNA triple helix at the 3' end (similar to a polyA tail) that

increases its stability. Additionally, both mRNAs and lncRNAs can be spliced, and some lncRNAs can be back-spliced and circularized (circular RNAs) (110).

LncRNA - Genomic location

One method of classifying lncRNAs is based on their location in the genome relative to protein-coding genes. Based on this genomic location of a lncRNA, the functional role of the lncRNA can be better predicted. As shown in Figure 6, lncRNAs can be defined as sense, antisense, bidirectional or divergent, intronic, intergenic or intervening(111, 112). Importantly, lncRNAs can overlap protein-coding genes and be expressed in a similar pattern as the coding gene(113) (Figure 6, IIa and IIb). The V3.0 lncRNA classifications by Arraystar Incorporated shows their analysis of the number of each subgroup. The majority of lncRNAs are classified as lincRNAs (19,590) followed by intronic (4,409), antisense (3,691), sense overlapping (1,597), and bidirectional (1,299) lncRNAs (www.arraystar.com).

Sense lncRNAs are transcribed in the same direction as exons of protein-coding genes (Figure 6, Ia) and have been found to be non-coding variants of genes that can regulate gene expression. A lncRNA can also partially overlap a protein-coding gene on the same genomic strand, and both are transcribed in the same direction (Figure 6, IIa). These sense-overlapping lncRNAs can be transcript variants of protein-coding genes, and the majority of these lncRNAs lack a substantial ORF for protein translation.

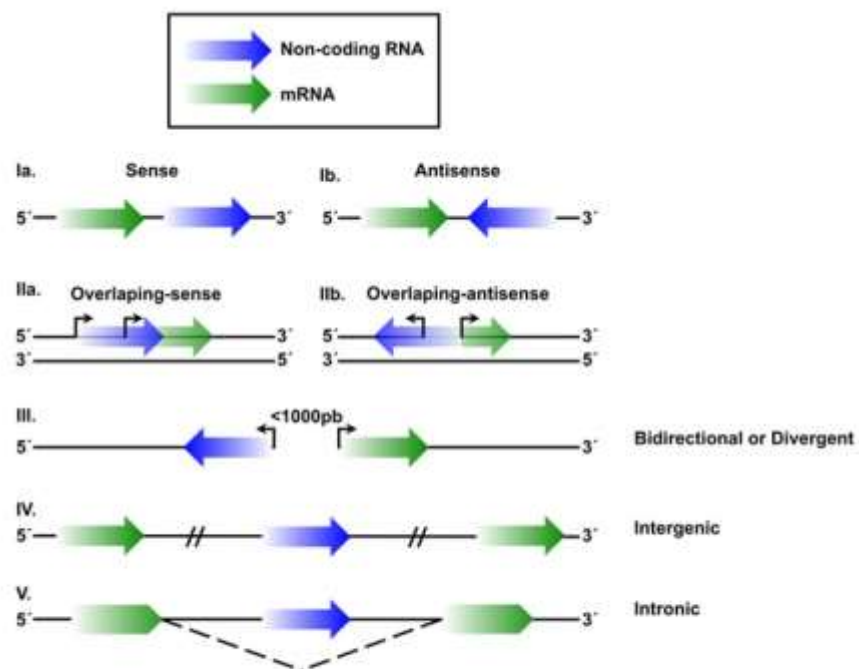


Figure 6: Classification of lncRNAs by genomic location.
Adapted from: Herrera-Solorio et al., *Clin Epigenetics*, 2017.

Bidirectional, or divergent, lncRNAs are transcribed in the opposite direction as a protein-coding gene within the vicinity (within 1kb) (Figure 6, III). There is similarity in expression patterns between some

lncRNA and its protein-coding counterpart such they may be regulated by similar mechanisms, but this is does not hold true for all bidirectional lncRNAs. Bidirectional expression of lncRNA-protein-coding gene pairs are prevalent in cancer.

Antisense lncRNAs are transcribed from the antisense strand, in the opposite direction as a protein-coding gene (Figure 6, Ib). A lncRNA can also partially overlap a protein-coding gene, and both are transcribed in the opposite direction (Figure 6, IIb). Antisense lncRNAs can share the same promoter as coding gene and their expressions can highly correlate(114). Antisense lncRNA may pair with a protein-coding gene on the sense strand to regulate epigenetic silencing, transcription, and mRNA stability(115). As an example of how genomic location can aid in characterization of lncRNAs, mouse sense-antisense transcripts tend to be poly(A) negative and localized in the nucleus(116). Antisense lncRNA transcription of genes affects both the expression and splicing of sense genes(117). Natural antisense transcripts' (NATs) expression regulate tumor suppressors and oncogenes in cancer. OncoNAT is useful tool that catalogs cancer-related genes with significant antisense transcription.

Intergenic lncRNAs (lincRNAs) are transcribed between two protein-coding genes and are at least 1kb away from the nearest protein-coding genes (Figure 6, IV). Very long intergenic ncRNAs (vlincRNAs) can range in size from 50 kilobases to 1 megabase. As of 2009, more than 1,000 lincRNAs were identified to be encoded by the mammalian genome and are clearly conserved across mammals. Some lincRNAs (e.g. HOTAIR) have been found to associate with chromatin modifying proteins and regulate gene expression(118). They have also been found to be important in a variety of cellular processes including apoptosis, cell-cycle progression, and cellular senescence. Intergenic lncRNAs can be subcategorized as being enhancer-associated (elncRNA) (present close to the promoter, in an enhancer region) or promoter-associated (plncRNA). ElncRNAs are more tissue-specific, less highly expressed, and less well conserved during evolution(119) and have been found to be important in certain diseases(120).

Intronic lncRNAs are transcribed from an intronic region of a protein-coding gene (Figure 6, V). Intronic lncRNAs can be transcribed in the sense or antisense direction. Intronic lncRNAs are expressed in similar tissues as the corresponding protein-coding gene, and the lncRNA may regulate alternative splicing or stability(121).

Additionally, there are certain regions of the genome that are associated with diseases, and these regions contain lncRNAs. Thus, these lncRNAs could be critical regulators in these diseases.

LncRNA - Excluding protein-coding potential

Classifying RNAs as non-coding and coding transcripts is not a straightforward process as mRNA and lncRNA exhibit similar structure, and a single RNA can function as an mRNA and also have a non-coding function. This discernment is also difficult, as lncRNAs overlap protein-coding genes (Figure 6, IIa and IIb). Certain lncRNAs have already been identified to encode functional peptides(8) (7), while some RNAs currently classified as lncRNAs may indeed encode proteins if researched further. These aforementioned difficulties of differentiating between coding and non-coding RNA are comprehensively reviewed by Dinger et al.(122).

Protein-coding transcripts contain a long open reading frame (ORF) capable of translating a protein, which can separate mRNAs from lncRNAs in some instances. Solely using ORF size is not acceptable as some RNAs classified functionally as lncRNAs also contain long ORFs (e.g. H19 and Xist). On the other hand, some mRNAs contain shorter ORFs. If an RNA contains an ORF, the conservation of that ORF can be analyzed to indirectly guide towards discerning between mRNA and lncRNA; certain codons for amino acids show preservation of protein-coding potential(111).

Certain RNAs classified as lncRNAs can contain ORFs by chance, and they might instead function solely as a ncRNA. On the other hand, if a lncRNA is identified to code for a protein it does not eliminate the possibility that it also has a critical non-coding function. Interestingly, there are specific bifunctional RNA that can have both protein-coding and coding-independent functions(123, 124). There are certain mRNAs that have regulatory non-coding RNA elements; for example, *oskar* RNA can act as a scaffold, independent of Oskar protein(125).

Research techniques (such as mass spectroscopy proteomics and *in vitro* translation assays) can be used to identify lncRNAs that encode functional peptides. Ribosome footprint profiling can be used to visually show the direct binding of transcripts to ribosomes, and this technique specifically can be used to analyze lncRNAs in mammalian cells(126, 127). Software tools, such as the Coding-Potential Assessment Tool (CPAT)(128), can be used to predict the likelihood that a transcript can encode for a protein by analyzing an ORF for the presence of known protein motifs.

LncRNA - Tissue specificity and evolutionary conservation

LncRNAs are typically tissue- and cell-type specific, even more than protein-coding genes(129-132). Since many lncRNAs are expressed in certain tissues, this could be such that they have critical functions in the development and maintenance of that specific tissue. Alternatively, in 2015, a group conducted a pan-cancer analysis of lncRNAs from multiple tumor tissues to their matched normal counterparts and found a class of lncRNAs that are altered in multiple cancers. They classified these as 'onco-lncRNAs' that are predicted to have conserved oncogenic and/or tumor suppressive across multiple cancer types(133).

Many lncRNAs have poor evolutionary conservation across species(134), but it is intriguing to characterize lncRNAs that are conserved and thus speculated to be critical for biological functions. Additionally, there are specific human lncRNAs that lack evolutionary conservation in non-human species, but are still functionally relevant in humans and should be studied further(134). LncRNAs have been identified in a variety of species, including yeast, viruses, and humans, but to further determine conservation of the lncRNA, samples from multiple species could be further analyzed for lncRNA expression. Lack of conservation does pose certain challenges in determining the research model for studying lncRNAs as many lncRNAs identified in non-human models have not progressed towards applicability in humans. LncRNAs can be spliced to produce multiple isoforms, which creates challenges in that a certain lncRNA could be spliced differently in a certain species, tissues, and under certain cellular stresses, thus making it difficult to determine the biologically functional isoform.

LncRNAs tend to exhibit poorer primary sequence conservation than mRNAs(134). Certain lncRNAs consist of highly repetitive sequences. The repetition of these elements can contribute to identification of function. For example, certain lncRNAs with Alu repeats can bind to hnRNPK and localize in the nucleus(135).

Secondary and tertiary structure are highly conserved and analyzing RNA structure could also help identify functionally important lncRNAs and their binding partners. LncRNAs are complex, high molecular weight molecules, which therefore makes crystallization difficult. The structure of an RNA can be determined by various *in vitro* and *in vivo* technologies, including dimethyl sulfate-sequencing (DMS-seq) and selective 2'-hydroxyl acylation analyzed by primer extension-sequencing (SHAPE-seq) (136). The presence of certain structural motifs can lead to the interaction with certain binding partners. For example, a G-rich motif in the lncRNA Braveheart interacts with certain nucleic acid binding partners such as a zinc-finger transcription factor (137).

LncRNA - Size, expression level, and stability

The size cut off of lncRNAs (>200 nucleotides) is an arbitrary cutoff that is used to separate from distinctly small ncRNAs. When comparing lncRNAs to mRNAs, lncRNAs tend to be shorter than mRNAs and have fewer but longer exons. Most lncRNAs range in size from less than 1kb (73%), between 1kb and 2kb (14%), and more than 2kb (13%)(138). Few lncRNAs are exceptions in that they can span 100 kilobases. For example, the promoter for the an antisense lncRNA Air (antisense Igf2r RNA) lies 107,796 base pairs away from the 3' end, in an intron flanking the Mas1 gene(139). To determine RNA size, Northern Blot analysis can be conducted.

There are a variety of cellular stresses that regulate lncRNA expression, including, but not limited to, DNA damage(140), viral infection(141), and serum starvation (quiescence)(142). LncRNA expression levels are dysregulated in many diseases and their expression can correlate with clinical parameters (e.g. overall survival). Many lncRNAs are expressed at low expression levels compared to protein-coding mRNAs(108). This poses research challenges and the debate about if lncRNAs are expressed at high enough levels to be functionally important. The stoichiometry of a lncRNA to its target gene can predict the likelihood of this interaction having a critical importance. To determine relative expression level of a lncRNA, quantitative reverse polymerase chain reaction (qRT-PCR) can be used, and if absolute expression is desired, *in vitro* transcription followed by qRT-PCR can be used.

In 2012, Clark et al. compared the stability of mouse lncRNAs and mRNAs. Based on half-lives, lncRNA stability is comparable to, but on average less than, mRNAs. Intergenic and antisense lncRNAs that act in cis are more stable than those derived from introns. Additionally, spliced lncRNAs are more stable than unspliced (single exon) lncRNAs. LncRNAs localized in the nucleus are more likely to be unstable(143). Comparing lncRNAs to mRNAs, mRNAs are primarily degraded by 5'-to-3' exonuclease digestion in the nucleus or decapping at their 5' end in the cytoplasm, while many unstable lncRNAs are degraded by the nuclear exosome or nonsense-mediated decay (NMD) in the cytoplasm(144).

LncRNA - Cellular localization

Comparing lncRNAs to mRNAs, mRNAs usually localize to ribosomes in the cytoplasm while lncRNAs can vary in their localization. LncRNAs are primarily localized in the nucleus but can also localize in the cytoplasm(131) (Figure 7). The localization of lncRNAs primarily in the nucleus is consistent with lncRNAs main function of contributing to epigenetic regulation of gene expression in the nucleus(129).

Many lncRNAs are exported from the nucleus to the cytoplasm through the canonical mRNA pathway, via the major export complex TREX and NXF1/NXT1 nuclear transport receptor(145).

The sequence of an RNA determines if it is a substrate for retention in the nucleus or exportation to the cytoplasm(145, 146). There are specialized cis-elements in RNAs that can promote retention in the nucleus. During splicing of an RNA, the spliceosome recruits export factors to the RNA that can promote exportation. There are splicing-independent factors that instigate nuclear export via association with the same exportation factors(145).

Certain lncRNAs are restricted to localization in the nucleus. For example, the most abundant and conserved lncRNA is MALAT1 (metastasis-associated lung adenocarcinoma transcript 1), also known as nuclear-enriched abundant transcripts 2 (NEAT2). NEAT1 is also restricted to the nucleus and both NEAT lncRNAs interact with nuclear speckles in human cells (147).

Common technique to determine lncRNA cellular localization are cellular fractionation and RNA fluorescence *in situ* hybridization (FISH).

Alterations in mRNA localization are associated with various diseases. Under cellular stresses, nucleocytoplasmic transport is disrupted by sequestering critical transport factors into stress granules to exacerbate neurodegeneration(148). Thus, it is considered that lncRNA mis-localization could also be related to diseases.

Knowledge of lncRNA localization provides insight into interactions and functions. Nuclear lncRNA can regulate transcription and chromatin remodeling in various ways, as shown in Figure 6. They can regulate transcription by acting as enhancer RNA (eRNA) (Figure 7A), or interacting with and recruiting chromatin modifying complexes (Figure 7B) or transcription factors (Figure 7C). lncRNA can also be the mediators to spatially bring chromosomes in close proximity (Figure 7D) and pre-mRNA splicing (Figure 7E). Cytoplasmic lncRNAs can regulate mRNA stability (Figure 7F) or mRNA translation (Figure 7G). Linear and circular RNAs can bind to and sponge miRNAs (Figure 7H). Some lncRNAs contain small ORFs that can be translated to small peptides (Figure 7I).

LncRNA -Biological function

Studying lncRNA biological function is critical, as lncRNAs already have been found to be associated with developmental processes (149) and dysregulation of lncRNAs is associated with diseases(150). Additionally, lncRNAs have been found to be involved with many biological processes, including proliferation, apoptosis, and differentiation by regulating gene expression, interacting with epigenetic modifiers, mRNA splicing and many others. LncRNA location can give a hint towards the function of a lncRNA, and some known nuclear and cytoplasmic functions are summarized in Figure 7. Broadly, lncRNAs molecular functions are subdivided into four categories: signals (HOTAIR), guides (Evl-2)(151), decoys (GAS5), and scaffolds (ANRIL) (Figure 3). LncRNA commonly form ribonucleoprotein (RNP) complexes to control the location of protein regulators (Figures 3 and 7).

LncRNAs can act in cis in the nucleus, where they modify chromatin and/or regulate nearby gene expression, or in trans, where they leave the site of their transcription to perform biological functions. It

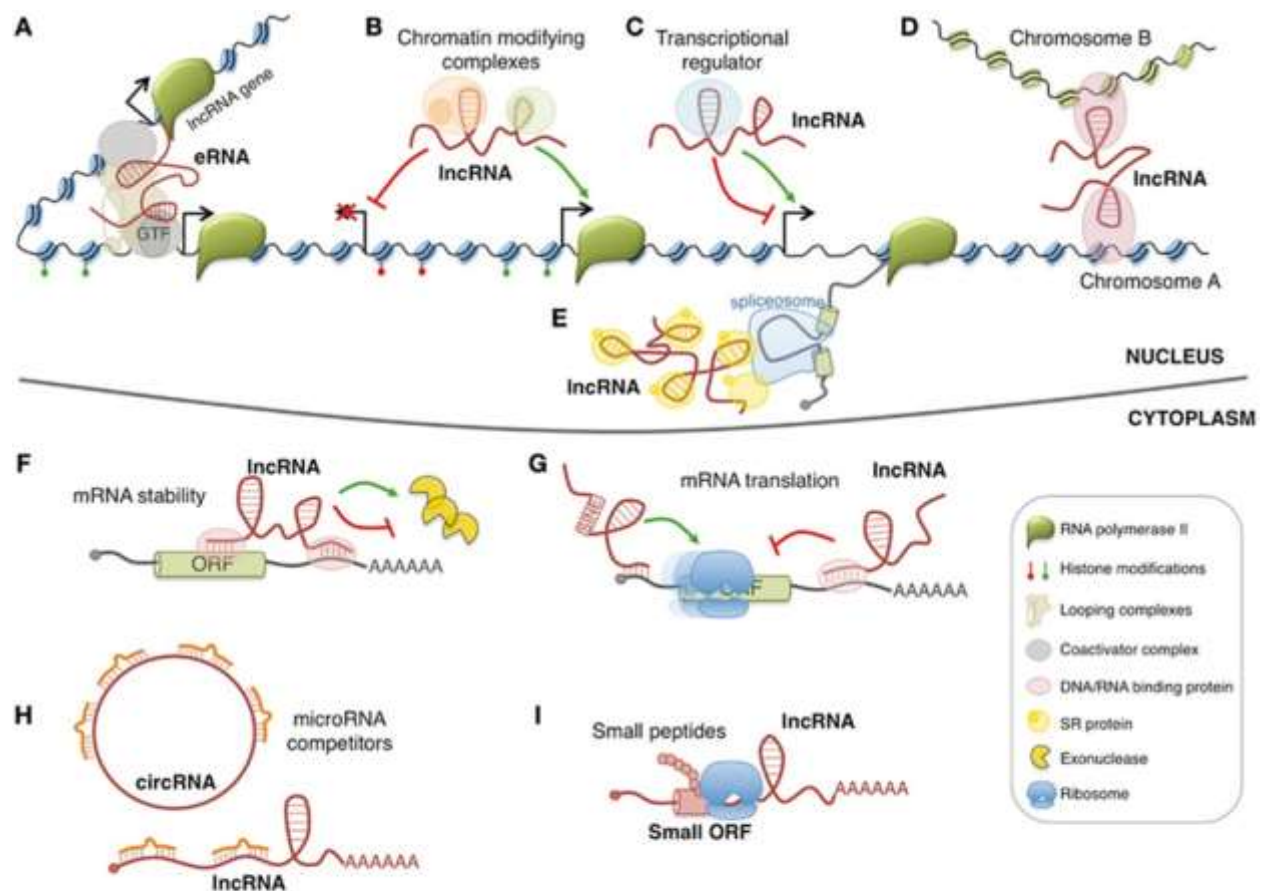


Figure 7: LncRNA functions in the nucleus and cytoplasm.
Adapted from: Morlando et al., *Mini Reviews in Medicine*, 2015.

can be difficult to determine the function of cis-acting lncRNAs as their expression can be directly, indirectly, or not important for downstream biological effects. The capability to act in cis makes them

unique from mRNAs, and allows lncRNAs to be involved in transcriptional regulation and chromosome looping.

Enhancer lncRNA

Figure 7A and 7D visually represent enhancer activity by a lncRNA. Recently, Miao et al. identified an enhancer lncRNA-mediated transcriptional regulation. The LEENE (lncRNA that enhances eNOS expression) lncRNA transcripts (located on chromosome 14) associates with the eNOS (endothelial nitric oxide synthase) locus (located in chromosome 7). Both chromosomes are brought into close proximity to permit the association and modulate eNOS mRNA synthesis(152).

Chromatin Remodeling

LncRNA can act as guides to recruit chromatin modifying enzymes to DNA to increase gene expression (Figures 7B and 3III). A recent study found that out of 3,300 human lincRNAs analyzed, 20% were bound by chromatin-modifying complexes including Polycomb Repressive Complex 2 (PRC2)(118). Certain lncRNAs can mediate chromatin remodeling in cis, while others can regulate in trans. In cis, transcribed lncRNAs such as *Air* can recruit the transcriptional repressive histone methyltransferase G9a to interact with the promoter of one of its imprinted genes *Slc22a3* and silence *Slc22a3* expression(153). In-trans, lncRNA such as HOTAIR can act as a scaffold and bind to key PRC2 components (Suz12 and Ezh2) and the LSD1-coREST complex to methylate gene histones and promote gene silencing(154).

Interaction with Transcription Factors

Nuclear lncRNA can bind to transcription factors (Figure 7C) and guide them to DNA (Figure 3III) or away from DNA (Figure 3II) to regulate transcription. LncRNAs have been found to activate transcription, while others suppress transcription. One example of a lncRNA contributing to transcriptional activation is *Evf-2* that is transcribed from an ultra-conserved enhancer, forms a stable complex with Dlx-2, which consequently activates Dlx-2 to enhance transcription of *Dlx-5/6(151)*. LncRNAs can repress transcription by binding to and sequestering transcription factors. An example of a lncRNA contributing to transcriptional repression is PANDA, which associates with the transcription factor NF-YA to act as decoy to keep NF-YA away from target gene chromatin. Due to this, pro-apoptotic genes are not expressed by NF-YA(155).

Splicing Regulation

As shown in Figure 7E, lncRNA can regulate splicing. More specifically, lncRNAs can regulate alternative splicing and interact with splicing factor as recently summarized (156). MALAT1 controls the activity of the SR protein family of splicing factors to regulate alternative splicing(157).

mRNA Stability

Figure 7F visually shows that lncRNA can regulate mRNA stability. For example, the lncRNA LAST (lncRNA-assisted stabilization of transcripts) stabilizes CCND1 mRNA. c-Myc regulates the lncRNA LAST, which in collaboration with CNBP binds to the 5' UTR of CCND1 mRNA to protect against possible targeting by nuclease(158).

mRNA Translation

As shown in Figure 7G, lncRNA can regulate mRNA translation. For example, lincRNA-p21 associates with JunB and β -catenin mRNAs and selectively decreases their translation by diminishing their polysomes (159). A lncRNA antisense to, and overlapping, the Uchl1 protein-coding gene was found to be required for Uchl1 translation. Furthermore, this was shown to be due to an embedded repetitive sequence (short interspersed nuclear element/SINE) in the lncRNA that regulates recruitment of polysomes to the mRNA for translation(115).

miRNA Regulation

As shown in Figure 7H, lncRNAs can interact with and sequester miRNAs and act as competing endogenous RNAs (ceRNAs). Multiple studies have analyzed for lncRNAs acting as miRNA sponges(160, 161). Targetscan (targetscan.org)(162) can be used to predict miRNA targets, while spongeScan (<http://spongescan.rc.ufl.edu>) (163) can be used to specifically identify miRNA binding elements in lncRNA sequences.

p53-regulated lncRNAs

Expression of the tumor suppressor p53 correlates with many host lncRNAs and these lncRNAs were induced in a p53-dependent manner(164). Alternatively, lncRNAs regulate p53 expression levels. For

example, the nuclear lncRNA PURPL (p53 upregulated regulator of p53 levels) associates with MYBBP1A (a protein that binds to and stabilizes p53) to decrease p53 stabilization, decrease p53 basal expression levels, and promote tumorigenicity in colorectal cancer(165).

Overall, lncRNAs lack a standard classification framework, which makes categorizing and analyzing lncRNAs difficult. Currently, in an attempt to standardize this process, lncRNAs are primarily classified by genomic location and function.

C. Circular ncRNAs

Circular ncRNAs - Overview

Circular RNAs (CircRNAs) are produced when precursor mRNA back-splice and the head to tail region bind each other. Due to their circular structure, they are thought to be stable as they would be resistant to degradation by exonucleases that typically degrade linear RNA. CircRNAs are typically cell- and tissue-specific, fairly conserved, and expressed at low levels(166). They are, however, expressed in high levels in brain, indicating they have neuronal functions. Additionally, circRNAs play a role in pathological conditions such as development of cancer and other diseases. For example, circRNAs are associated with EMT-related functions(167).

Circular ncRNAs - Functions

The functions of circRNAs are currently being researched. Specific techniques are required to identify circRNAs, such that large scale high-throughput sequencing of ncRNAs looking for linear, polyA+ RNAs would not detect circRNA expression. When desired, circRNA detection efficiency of RNA-seq with a specific microarray platform could be used to efficiently profile circRNAs by circRNA microarray. Using a specific microarray platform, thousands of circRNAs could be detected in cell-free plasma samples(168). Some known functions are sequestering miRNAs and proteins, modulating transcription, interfering with splicing, and translating proteins (Figure 8).

CircRNAs can function as miRNA sponges, thus preventing the miRNA from targeting mRNA and permitting translation (Figure 8A). One of the most well-known circRNAs is ciRS-7 (also known as CDR1as), which contains 70 binding sites for miR-7 and has been identified as a miRNA inhibitor. Consequently, this regulation is involved with many different cancers since miR-7 directly targets several

oncogenes(169). Additionally, many other circRNAs acts as miRNA sponges. A testis-specific circRNA known as sex-determining region Y (circSry) serves as a miR-138 sponge to regulate testes development (170).

Exon-intron circRNAs (EircRNAs) can enhance transcription of their parental genes. For example, after EircRNA formation, it can interact with RNA Pol II and U1 snRNP to enhance transcription (Figure 8B).

CircRNAs can also act as a protein scaffold (Figure 8C) or bind and sequester proteins (Figure 8D) to particular cellular locations to modulate protein interactions(171). For example, circMbl can bind to the muscleblind (MBL) protein to regulate MBL function(172).

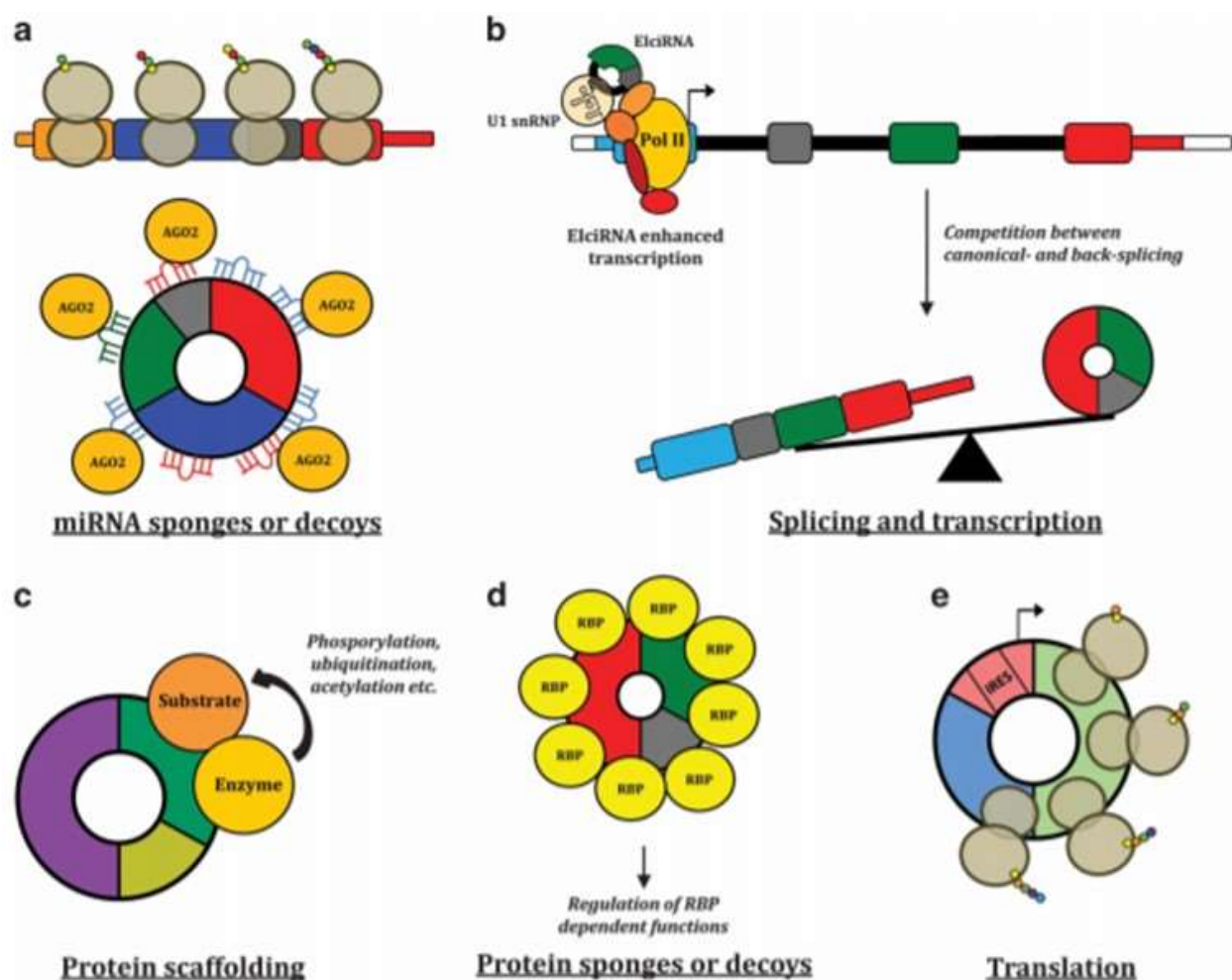


Figure 8: Circular RNA functions.

Adapted from: Kristensen et al., *Oncogene*, 2018.

Hundreds of endogenous circRNAs containing an internal ribosome entry site (IRES) and AUG start codon can interact with ribosomes and have translation abilities(173) (Figure 8E). Using human embryonic kidney 293 cells and HPV-18 positive cervical cancer cells (HeLa), a group identified dependence of N6-methyladenosine (m⁶A) motifs is the sequence of these circRNA molecules capable of translating proteins(174).

As host cells produced circRNAs, certain viruses also produce circRNAs. For example, Epstein Barr virus (EBV) encodes latent and lytic viral circRNAs, some of which are dysregulated in stomach cancer samples(175). Kaposi's sarcoma herpesvirus also can encode viral circRNAs that are up-regulated during the lytic phase of infection and alter cell growth(176).

D. LncRNAs - Clinical Implications

Currently, cervical cancer screening is the combination of high-risk HPV DNA testing with cytological morphological assessment during Pap smear. In the case of cervical cancer, identification of biomarkers associated with oncogenic progression would be useful. It is proposed by some that panels of epigenetic biomarkers (e.g. high CpG methylation rates of DAPK(177), CDH1(178), CCNA1(179), and CADM1(180)) analyzed during cervical cancer screening could be more beneficial predictors of oncogenic progression than standard methods. Certain epigenetic markers suggested are expressed in plasma or serum and can detect stages less invasively (DAPK(177) and CDH1(178)), while others are expressed in the tissue specimen to distinguish between stage and duration of disease (CCNA1(179) and CADM1(180)). Epigenetic alterations related to ncRNAs (including lncRNAs) contribute to cervical cancer and are being evaluated to determine disease progression and potentially target them for therapeutic benefit(181).

LncRNAs are strikingly more tissue specific than protein coding genes(129-132). Clinically, this allows for the advantage that these specific lncRNAs can be targeted for therapy to alter solely the tissue it is expressed in to reduce off-target effects. LncRNAs could be potentially targeted by multiple methods, including RNAi (siRNA), antisense oligonucleotides (ASOs), and small molecules(182, 183). Additionally, lncRNAs are being researched as clinical biomarkers for cancer detection and prognosis.

There are various basic science and clinical research studies being conducted to analyze for clinical potential of lncRNAs. For example, inhibiting the oncogenic lncRNA HOTAIR *in vitro* can inhibit cancer invasiveness, especially in cells with high PRC2 activity(184). Certain groups are working to identify a

gene signature to predict prognosis of cancer patients. One group specifically identified a prognostic 15-lncRNA signature for cervical squamous cell carcinoma (CSCC) that are associated with patient survival(185). There is a Phase 2/3 clinical trial designed to validate a previously identified and independently validated prognostic and predictive mRNA-lncRNA signature for triple-negative breast cancer (TNBC), which could be used to classify TNBC patients into high- or low-risk of recurrence. This signature was used to identify high-risk TNBC patients, such that this decision could influence the decision of alternative therapeutic treatments on this set of high-risk patients compared to low-risk. This study is actively recruiting patients (NCT02641847). A recently completed clinical trial is analyzing expression of various ncRNAs (lncRNAs, snRNAs, snoRNAs, miRNAs) in radiation-induced fibrosis in breast cancer patients. In addition to ncRNA expression, they are also analyzing for alterations pre- and post-radiation in cellular localization. This study was completed in 2018 (NCT03000764).

The idea of collecting a patient's body fluid (e.g. urine and blood) to analyze for expression of circulating lncRNA(s) has become intriguing in the field. It could eliminate the current invasive diagnostic and prognostic method to biopsy patient tissue. As examples, lncRNA PVT1 is a potential serum biomarker for detection of cervical cancer(186), and GAS5 is proposed as a diagnostic biomarker in non-small cell lung cancer(187). Excitingly, detection of lncRNAs in body fluids of patients *has* translated to being valuable to detect cancer in human patients. The lncRNA prostate cancer antigen 3 (PCA3) is leading the field of lncRNAs commercially available for cancer detection. The urine test PROGENSA PCA3 was FDA-approved in 2012 for use in men who are considering repeat biopsy after an initially negative result for prostate cancer. It is considered to be more specific and effective in identifying prostate cancer than the prostate-specific antigen (PSA) urine test. Certain clinical trials being conducted, in addition to their primary objective, are collecting blood/serum samples at diagnosis and later stages (e.g. at time of surgery, pre- and post-surgery) to detect alterations in lncRNA biomarkers. For example, a Phase 2/3 clinical trial to treat hormone receptor positive and triple-negative breast cancer (TNBC) is additionally collecting samples to test patients for angiogenic serum markers (miRNA, lncRNA, and circRNA) measured at diagnosis and time of surgery. This study is active, but currently not recruiting patients (NCT02221999). HOTAIR, as a single oncogenic lncRNA, expression is monitored in the peripheral blood of thyroid cancer patients. This study is not yet recruiting patients (NCT03469544). Analysis of ncRNA expression in exosomes is currently an actively pursued direction of clinical research. One newly posted clinical trial is currently recruiting patients to analyze ncRNA expression via genetic sequencing of miRNA/lncRNA in the exosome of epithelia ovarian cancer patients to determine biomarkers for detection and prognosis of patients. This study is currently recruiting patients(NCT03738319). Unfortunately, there are no results posted on clinicaltrials.gov for any lncRNA studies at this time.

IV. Human Papillomavirus (HPV)

A. HPV - Overview

Human papillomavirus (HPV) is now recognized as one of the most prevalent sexually transmitted infection in the world. It is a member of the *Papillomaviridae* family, which is a group of small, non-enveloped viruses with double stranded DNA circular genomes that are ~8 kilobase pairs. More specifically, HPV is classified as a mucosal alpha-virus. In the 1980s, Harald zur Hausen's group discovered that cervical carcinoma samples contained DNA from certain strains of HPV, which suggested HPV as a etiological factor in cervical cancer(188). A schematic of the HPV-mediated progression to carcinogenesis is presented in Figure 9. Upon micro-abrasions in cutaneous and mucosal sites, the basal layer of stratified squamous epithelia is exposed and the virus can infect keratinocytes. Infected cells in this region have a highly organized process of tissue renewal that viruses utilize to replicate and produce virions to infect other cells. In early stages of HPV infection, the viral genome is maintained in an episomal state; as the infected cells differentiate, the viral genome replicates in the nucleus of the infected cells using the host cells replication machinery. As the virus-infected cells differentiate and progress towards the surface of the epithelium, the virus replicates alongside the host cell in a differentiation-dependent manner. Once the surface of an infected cell reaches the surface of the epithelium, virions are assembled and released to infect other target cells.

There are more than 100 HPV strains, which are divided into the two subtypes high-risk and low-risk, depending on the clinical prognosis of lesions after HPV infection. Low-risk HPVs (e.g. HPV-6 and -11) can cause benign epithelial lesions and

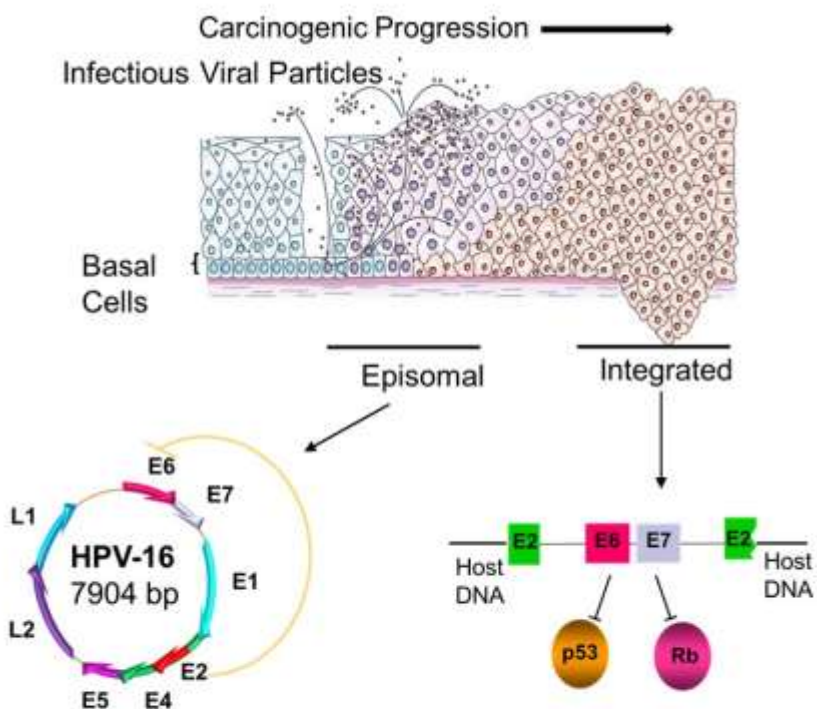


Figure 9: HPV-mediated progression to cancer.
Adapted from: Woodman et al., *Nature Reviews Cancer*, 2007.

are associated with 90% of genital warts. There are fifteen strains associated with cancer, and these are classified as high-risk HPVs (e.g. HPV-16 and -18). Each of these high-risk strains encode for 6 genes in the early region (E1, E2, E4, E5, E6, E7) and two genes in the late region (L1 and L2). The genes encoded by the early region are responsible for modulating the cellular environment for viral replication and immune evasion, while the genes from the late region produce capsid, structural proteins(189).

As stated above, the viral genome is maintained in an episomal state in early HPV infection. While the genome is episomal, the E2 gene can repress the expression of the E6 and E7 oncogenes. However, upon persistent infection, over many years, the virus can integrate into the host genome. When integration occurs, the E2 gene is disrupted such that it cannot repress the promoter of E6 and E7, and there is a robust increase in the expression of these oncogenes. Up-regulation of E6 and E7 adds to carcinogenic progression because they contribute to the degradation of two important tumor suppressors p53 and Rb, respectively(189) (Figure 9). The co-expression of the oncogenes immortalizes the cells(190). After many years of persistent HPV infection, other mutations in host genes occur that transform the cells and contribute to formation of a tumor(191). In addition to viral infection, genetic mutations as well as epigenetic alterations contribute to carcinogenesis(181).

HPV infections are asymptomatic and usually cure spontaneously by the immune system; however, persistent high-risk infection is an etiological agent in a variety of different cancers. In the United States, there are more than 42,000 HPV-associated cancers each year related to cervical (91%), vaginal (75%), vulvar (69%), penile (63%), anal (91%) and oropharynx (70%) cancers in the United States. In 2012, worldwide, 4.5% of all cancers worldwide (1.2 million) are related to HPV infection, with 73% of these cases related to the two strains HPV-16 and -18 combined(192).

There are currently three FDA-approved prophylactic vaccinations against HPV infection. Gardasil 9 protects against HPV-16 and -18, low-risk HPV-6 and -11, as well as five additional high-risk strains HPV-31, -33, -45, -52, and -58. Gardasil 9 has now replaced Gardasil (protects against HPV-16 and -18, as well as low-risk strains HPV-6 and -11). Cervarix protects against the two most carcinogenic strains (HPV-16 and -18). A review published in 2017 summarizes these prophylactic vaccines in more detail, describing that Gardasil and Cervarix approach 90% effectiveness in preventing HPV-16/-18 infections in women older than 25 years(193).

B. HPV - Viral Oncoproteins

carcinogenesis. E6 interacts with many proteins to activate transcription of telomerase reverse transcription (TERT); combined with Rb inactivation by E7, this is an essential step to immortalization. E6 also mediates degradation of PDZ proteins, which leads to loss of cell polarity and induces hyperplasia.

Some known downstream effectors of HPV E7 are shown in Figure 11. E7 is most well-known for being involved in the regulation of retinoblastoma (Rb) family of proteins, contributing to cell cycle activation and proliferation downstream. E7 mediates pRB and pRB-related proteins (p107 and p130) destabilization by a proteasome-dependent mechanism. When E7 binds Rb, this dissociates the transcription factor E2F-1 that is bound to Rb such that unbound E2F-1 is free to drive expression of S phase genes and induce cell cycle activation and proliferation. E7 abrogation of Rb function leads to increased expression of the tumor suppressor p53. Consequently, E6 has evolved to target p53. E7 can also regulate many other proteins to influence genomic instability, apoptosis, and proliferation and contribute to carcinogenesis.

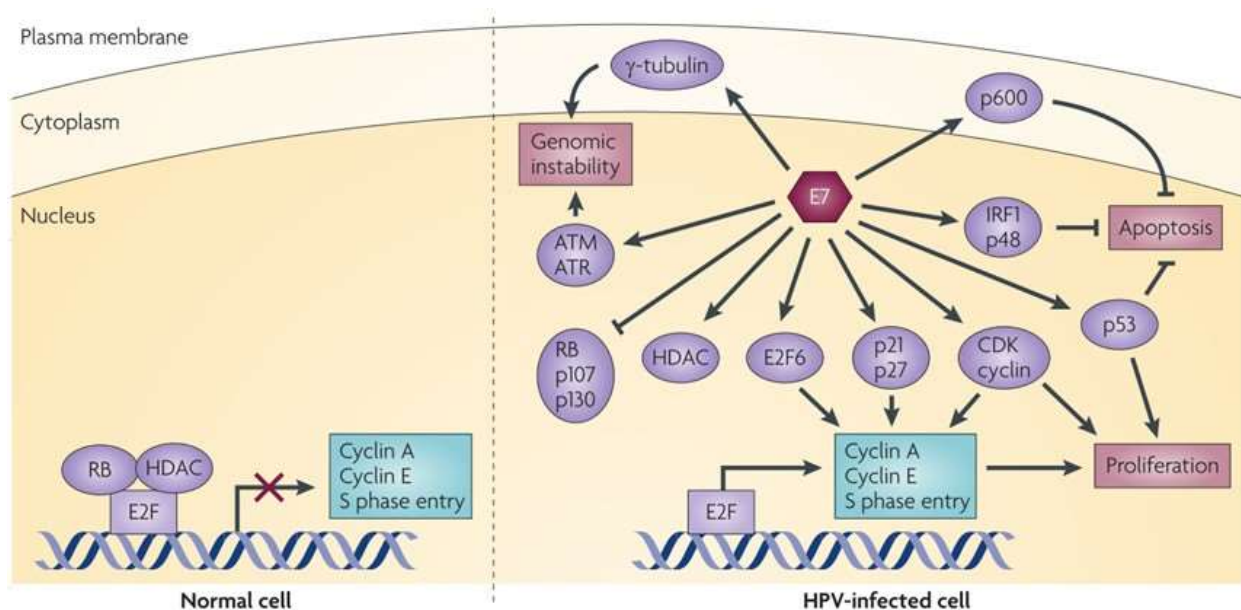


Figure 11: HPV E7 oncoprotein regulation

Extracted from: Moody et al., *Nature Reviews Cancer*. 2010.

V. LncRNAs in cervical cancer

A. Cervical Cancer - Overview

Cervical cancer is one of the most common gynecological cancers. Worldwide, it is responsible for the second highest number of deaths in female cancers (exceeded only by breast cancer), as 10-15% of all

female cancer-related deaths were due to cervical cancer(205). 80% of patients have already developed invasive cancer at the time of diagnosis.

A combination of radiation therapy and low-dose chemotherapy is often used to treat early stages of cervical cancer. Alternatively, radiation therapy alone or surgery could be used for a small, early-stage tumor. Patients at high risk for recurrence could be administered radiation therapy and chemotherapy after surgery. Recurrent cervical cancer or patients with metastasis will be administered a combination therapy of platinum-based chemotherapy with the targeted therapy bevacizumab (Avastin) (<https://www.cancer.net/cancer-types/cervical-cancer/treatment-options>).

Identifying biomarkers for early detection as well as for prognosis are desirable. Somatic mutations in PIK3CA, PTEN, TP53, and KRAS are involved with the pathogenesis of cervical cancer(205). In the future, to predict, diagnose, and treat cervical cancer, it is proposed that utilizing lncRNAs could be beneficial.

B. Host lncRNAs are altered in cervical cancer

Unfortunately, The Cancer Genome Atlas (TCGA) does not directly include all of the human lncRNAs. However, due to the identification of lncRNAs role in cancer, many databases have been developed to summarize specific lncRNAs expression correlated with survival. MD Anderson developed TANRIC (The Atlas of non-coding RNA in Cancer), which extracted RNA sequencing data from TCGA and provided lncRNA expression correlative to survival in cancer patients. Currently, there are 307 cases of cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) available (<https://portal.gdc.cancer.gov/projects/TCGA-CESC>).

There have been multiple analyses of global differential expression of host lncRNAs in cervical cancer that have assisted in identifying specific lncRNAs playing a role in cervical cancer. Interestingly, a single lncRNA could be oncogenic in one tissue type, but tumor suppressive in another tissue lineage. Expression changes of lncRNAs in cancer versus normal samples as well as how altering lncRNA expression affects cancer phenotype (e.g. proliferation, migration, invasion, etc.) are useful; however, further elucidation of specific mechanisms of their regulation is critical for applicability and clinical potential from these lncRNA studies.

Host lncRNAs are altered at the three pre-malignant stages of cervical dysplasia as well as in established cervical cancer samples compared to normal samples. One group conducted differential lncRNA expression profile analysis between mild (CIN1), moderate (CIN2), and severe (CIN3) cervical dysplasia compared to non-neoplastic cervical tissue. Out of the 1,056 lncRNAs analyzed, 13 were aberrantly expressed in all three CIN stages(206).RNA-seq was performed on HPV-16 positive cervical cancer samples and matched adjacent non-tumor samples to identify differential expression of lncRNAs, circRNAs, miRNAs, and mRNAs in human patients, including 3 novel lncRNAs and 44 novel circRNAs. Additionally, this study analyzed the ceRNA network to show that each miRNA targeted multiple lncRNAs and circRNAs(207). In HPV-18 positive cervical cancer samples, the differential expression of circular RNAs with and without radiation was analyzed by RNA-seq. Further, the mechanistic role of these circular RNAs was evaluated in radioresistance. Circular RNA functions and pathways were analyzed, and more specifically circRNA-miRNA target gene interaction network studies were conducted(208).A recent study provided 41 microarray datasets and 10 RNA-seq data sets that can be analyzed for circRNA expression in cervical cancer(168). Overall, this provides resources that show host lncRNA alterations, as well as provides resources to extract host lncRNAs potentially involved in cervical cancer initiation and progression.

Many specific lncRNAs have been found to be involved in cervical cancer, some of which are HOTAIR, MALAT1, CCHE1, and lncRNA-EBIC(209, 210) and a summary of cervical cancer-related lncRNAs can be seen in Figure 12 below.

HOTAIR (Hox transcript antisense RNA)

HOTAIR is one of the most studied lncRNAs in cervical cancer. It is an oncogenic lincRNA that, in comparison to normal tissue, is up-regulated in many human cancers, including breast, gastric, colorectal, and cervical cancer(211-213). It interacts with the PRC2 complex leads to epigenetic regulation of many different cellular pathways(213).

In the context of cervical cancer, increased HOTAIR expression correlates with lymph node metastasis(211) and tumor size(211), tumor stage(212), and depth of cervical invasion(212). HOTAIR expression has been shown as a prognostic factor as high HOTAIR expression significantly correlated with poor overall survival(212) and disease-free survival/recurrence (211, 212) in cervical cancer patients. Knockdown of HOTAIR expression reduced cell proliferation, migration, and invasion in cervical cancer cell lines. HOTAIR could be exerting these oncogenic effects by regulating vascular endothelial growth

factor (VEGF), matrix metalloproteinase-9 (MMP-9), and epithelial-to-mesenchymal transition (EMT)-related genes(211) (Figure 12a). HOTAIR also enhances radio-resistance by inhibiting p21 in cervical cancer(214). Thus, HOTAIR could be used as a prognostic marker or therapeutic target in cervical cancer.

MALAT1 (Metastasis-associated lung adenocarcinoma transcript 1)

LncRNA MALAT1 [also known as nuclear-enriched transcript 2 (NEAT2)] is dysregulated in many cancers, including lung(215) and cervical cancer(216). Depending on the cancer type, it can be oncogenic or tumor suppressive.

In the context of cervical cancer, MALAT1 is up-regulated in tumor tissues compared to adjacent normal tissue, and this high expression correlates with tumor size, stage, invasion, lymph node metastasis, and poor overall survival(217). MALAT1 is involved with cervical cancer cell proliferation, cell cycle, and invasion and can regulate gene expression of apoptotic regulations (e.g. Bcl-2 and caspase-3)(216). Furthermore, MALAT1 regulates cancer cell growth and invasion at least in part by regulating miR-124 and its downstream target RBG2(218) (Figure 12f). MALAT1 could be a diagnostic or prognostic marker in cervical cancer.

LncRNA-EBIC (EZH2-binding lncRNA in cervical cancer)

Sun et al. conducted a microarray analysis of lncRNA and mRNA differentially expressed in cervical cancer and paired peritumoral tissues. They identified 708 up- and 836 down-regulated lncRNAs, including the up-regulated lncRNA EBIC (also known as lncRNA-TI17313). This lncRNA plays a role in migration and invasion in cervical cancer cells. Additionally, lncRNA-EBIC binds to EZH2 to regulate transcription in cervical cancer, and this interaction is required for the repression of E-cadherin (tumor suppressor that inhibits epithelial to mesenchymal transition (EMT) and prevent malignant progression)(219) (Figure 12e).

ANRIL (Antisense non-coding RNA in the INK4 locus)

ANRIL lncRNA plays an oncogenic role in many cancers, including melanoma, osteosarcoma, and cervical cancer(220). High ANRIL expression in cervical cancer patients correlated with tumor, lymph node metastasis, and poor overall survival(220). ANRIL plays a role in cell proliferation, migration, and invasion

in cervical cancer(220) (Figure 12g). ANRIL can also promote cervical cancer developing by sponging miR-186(221).

HOXA11-AS (Homeobox A11 antisense)

Microarray analysis has been used to identify differential expression of lncRNAs in human cervical cancer tissue samples(222, 223). A specific microarray analysis of mRNAs and lncRNAs expression identified differential lncRNA expression, number of differentially up- or down-regulated lncRNAs and mRNAs in each chromosome, and types (sense, antisense, intronic, intergenic, or bidirectional) of lncRNAs up- and down-regulated in cervical cancer(222). A specific microarray analysis identified HOXA11-AS lncRNA alteration in cervical cancer compared to normal cervix. Further analysis of this lncRNAs showed that HOXA11-AS is a regulator of HOXA11 that contributes to cervical cancer(223).

SNHG1 (Small nucleolar host gene 1)

Another group sent three pairs of cervical cancer tissue samples and corresponding adjacent normal samples for lncRNA microarray identify differential expression of lncRNAs. From this analysis, they identified that lncRNA SNHG1 is up-regulated in cervical cancer, and knockdown of this lncRNA resulted in a decrease in cell proliferation, migration, and invasion in both HPV-positive and HPV-negative cervical cancer cell lines(224). A recent review summarized that SNHG1 expression is up-regulated in at least 11 types of cancers and contributes to proliferation, migration, and invasion of cancer cells; therefore, it is proposed to consider SNHG1 as a prognostic marker and/or therapeutic target(225).

LncRNA CCHE1(Cervical carcinoma high-expressed 1)

In cervical cancer, increased CCHE1 expression is associated with tumor size, stage, and poor survival(226, 227). CCHE1 is involved with proliferation and physically associates with proliferating cell nuclear antigen (PCNA) mRNA to enhance the PCNA expression (Figure 12h). Thus, CCHE1 could potentially serve as a diagnostic or prognostic marker in cervical cancer(227).

LncRNA CRNDE (Colorectal neoplasia differentially expressed RNA)

LncRNA CRNDE (colorectal neoplasia differentially expressed RNA) is altered many cancers, including colorectal cancer(228), glioma(229) and expression correlates with patient prognosis(230). Up-

regulation of CRNDE expression in cervical cancer was determined from a lncRNA microarray analysis of cervical cancer tissues compared to normal adjacent tissues. Knockdown of CRNDE in HPV-positive and -negative cervical cancer cell lines decreased cell proliferation, migration, and invasion(231). Recently, CRNDE was found to be associated with poor prognosis in cervical cancer patients, and it targets PI3K/AKT to promote cervical cancer cell proliferation and inhibition of apoptosis(232).

MEG 3 (Maternally expressed gene 3)

MEG3 is a tumor suppressor that mediates p53 signaling and inhibits cell proliferation. It has been found to be down-regulated in a variety of human cancers, including cervical. The MEG3 is a tumor suppressor in cervical cancer, as it inhibits cervical cancer cell proliferation through induction of cell cycle arrest and apoptosis(233) by regulating miR-21-5p(234) (Figure 12i).

BCAR4 (Breast cancer anti-estrogen resistance 4)

The lncRNA BCAR4 has been shown to enhance cell proliferation and promote metastasis in breast cancer by activating the HER2/3 pathway(235). Due to this, in cervical cancer, lapatinib (EGFR/HER2

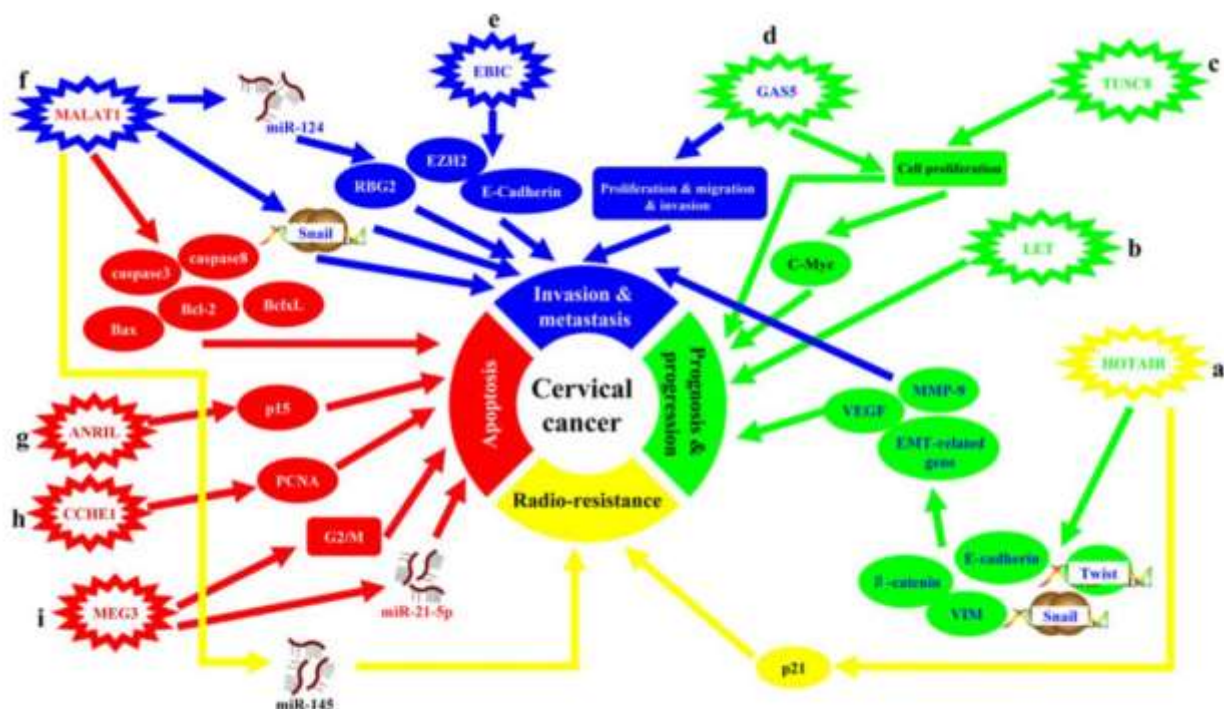


Figure 12: LncRNA functions in cervical cancer.
Extracted from: Peng et al., *Tumor Biology*, 2016.

inhibitor) BCAR4 lncRNA has been found to be associated with response to lapatinib in human patients(236).

LncRNA Loc554202

LncRNA Loc554202 is the host gene of miR-31(237). It has been shown to be involved in the development of breast(238), colorectal(239), and cervical cancer(240). Loc554202 exhibits higher expression in cervical cancer tissues than adjacent non-cancer tissues. High expression of the lncRNA correlates with poor overall survival, tumor size, stage, and lymph node metastasis. Knockdown of Loc554202 inhibits cell proliferation and induces apoptosis in cervical cancer cell lines(240).

H19

H19 can be oncogenic(241) or tumor suppressive (242), depending on the tissue lineage. In cervical cancer, it has been shown to be oncogenic and regulate cell proliferation(243). Additionally, in cervical cancer, H19 can sponge miR-138-5p to regulate proliferation and apoptosis(244).

GAS5 (growth arrest-specific transcript 5)

GAS5 is down-regulated in cervical cancer tissues as well as many other cancer types (e.g. breast(245) and lung(246))(247). GAS5 is a tumor suppressor lncRNA that when knocked down *in vitro* led to increased cell proliferation, migration, and invasion in cervical cancer (Figure 12d). The decrease in GAS5 expression significantly correlated with advanced cancer progression(248), thus GAS5 is considered a potentially useful biomarker for cervical cancer patients.

TUSC8 (tumor suppressor candidate 8)

Another tumor suppressor lincRNA in cervical cancer is TUSC8 (also known as XLOC_010588). In cervical cancer, TUSC8 expression is dramatically reduced, and this lower expression correlates with tumor stage and tumor size. With low TUSC8 expression in cervical cancer, c-myc is up-regulated and cervical cancer cells exhibit increased proliferation. Mechanistically, TUSC8 interacts with c-myc to consequently down-regulate c-Myc expression, and exogenous expression of TUSC8 in cervical cancer cells revealed decreased c-myc expression and decreased proliferation(249) (Figure 12c). Therefore, TUSC8 could be a potential therapeutic target in cervical cancer.

LncRNA LET

LncRNA LET is down-regulated in cervical cancer, as well as many other cancers (colorectal, hepatocellular, and lung), and low expression correlates with poor overall survival. Decreased expression of lncRNA LET correlates with tumor stage and lymph node metastasis(250) such that it might be a useful prognostic marker and therapeutic target for cervical cancer (Figure 12b).

AC017078.1 and XLOC_011152

Recently, a group analyzed systemic lncRNA expression in serum from cervical cancer patients via microarray analysis. Further validation discovered two potential candidate lncRNA biomarkers whose expression was down-regulated in cervical cancer (AC017078.1 and XLOC_011152)(251). Further elucidation of circulating lncRNAs that could be used as blood-based biomarkers in cervical cancer diagnosis could be beneficial.

hsa_circ_0018289

Microarray analysis was used to identify differentially expressed circRNAs in cervical cancer tissues and reveal up-regulation of hsa_circ_0018289. Further analysis of this specific lncRNA showed its role in proliferation, migration, and invasion of cervical cancer cells as well as binding and sponging miR-497 to contribute to cervical cancer(252).

hsa_circ_0023404

Hsa_circ_0023404 was shown to be up-regulated in cervical cancer tissue compared to normal tissue, and high expression of this circRNA correlated with poor overall survival. Knockdown of hsa_circ_0023404 reduced proliferation, migration, and invasion in cervical cancer cells. This circRNA binds to and sponges miR-136 such that the miRNA cannot regulate its target TFCEP2, which consequently leads to activation of YAP signaling and cervical cancer progression(253).

circRNA-000284

Another group conducted microarray analysis of cervical cancer cells (HPV-16 positive, HPV-18 positive, and HPV-negative) compared to normal cervical cell line to identify differential circRNA expressions and more specifically identify circRNA-000284 as significantly up-regulated in all cervical cancer cell lines analyzed. This specific circRNA binds to and sponges miR-506 to promote proliferation and invasion in cervical cancer(254).

CircRNA8924

Human circRNA microarray V2.0 was used to identify up-regulation of circRNA8924 expression in cervical cancer tissues. Further studies showed the involvement of the circRNA in proliferation, invasion, and migration by competitively binding to miR-518d-5p/miR-519a-5p family to modulate CBX8 expression in cervical cancer(255).

CircPAIP2

An intron containing circRNA (EIciRNA) circPAIP2 was shown to regulates memory-related genes through PABP reactivation in human cervical cancer cells(256).

VI. LncRNAs regulated by HPV

A. HPV-regulated host lncRNAs

Many host lncRNAs are dysregulated upon HPV infection with the entire HPV genome. A 2016 review summarized lncRNAs that were altered upon HPV status, including many well-studied lncRNAs (e.g. XIST, H19, and MEG3) (DOI: 10.4137/ATV.S29816). CDKN2B-AS, EGOT (Eosinophil granule ontogeny transcript), NCRNA00185, PRINS (Psoriasis Susceptibility-related RNA Gene Induced by Stress), TTTY14, TTTY15, XIST, and LINC00152 were all identified to have altered expression in HPV active versus HPV negative oral tumor samples(257). LncRNA CCHE1 expression is correlated with HPV status in cervical cancer(226). Microarray analysis of HPV-16 positive cervical cancer cells compared to HPV-negative cervical cancer cells was conducted and 4750 differentially expressed lncRNAs were identified(258). RNA-seq analysis of HPV-16 positive tumor tissues compared to normal adjacent tissues identified lncRNAs and circRNAs that are differentially expressed(207). Nohata et al. conducted RNA-seq analysis of HPV-positive versus HPV-negative HNSCC tumors and identified 140 lncRNAs significantly differentially expressed between the groups. Multiple other molecular changes occur between HPV-positive

and -negative HNSCC samples, however, it can be speculated that HPV is contributing to alterations in lncRNA expression profiles(259).

Additionally, many host lncRNAs are regulated specifically by the sole presence of both the HPV E6 and E7 oncoproteins. MEG3, PCNA-AS, and H19 were all differentially expressed in human foreskin keratinocytes expressing HPV-16E6/E7(260). HPV proteins can alter mitochondrial lncRNA expression to potentially regulate the cell cycle. More specifically, HPV proteins induces the expression of a sense mitochondrial ncRNA (SncmtRNA-2)(261).

Certain lncRNAs can be altered solely by the expression of the HPV E6 oncoprotein. For example, CCEPR (Cervical carcinoma expressed PCNA regulatory) lncRNA expression is regulated by HPV E6 oncoprotein to alter cell proliferation(262). Differential expression of lncRNAs was determined in human foreskin keratinocytes with and without HPV-16 E6 expression. FAM83H-AS1 expression is regulated by HPV-16 E6 oncoprotein in a p53-independent, p300-dependent manner. FAM83H-AS1 up-regulation altered cell proliferation and migration in HPV-16 positive cervical cancer cells (Chapter 2).

Alternatively, certain lncRNAs can be altered by the sole expression of the HPV E7 oncoprotein. A group conducted a study in the HPV-18 positive HeLa cervical cancer cell line where they knockdown HPV-18 E7 with siRNA and conducted lncRNA microarray analysis. They identified LINC01101 and LINC00277 whose expression increased with HPV-18 E7 knockdown. Both of these lncRNAs exhibited decreased expression in pre-cancerous lesion and cervical cancer compared to control tissues in a patient cohort of HPV-16 and -18 positive status. Additionally, lncRNA LINC01101 and LINC00277 expressions were significantly decreased in high-risk HPV-positive samples compared to HPV-negative samples; these two lncRNAs are more drastically decreased in HPV-18 samples than HPV-16, but both follow the same trend(263). CircRNA expression changes altered by the high-risk HPV-16 E7 oncoprotein were analyzed in a cervical cancer cell line (CaSki). More specifically, CaSki cells with knockdown of endogenous HPV-16 E7 underwent microarray analysis to identify 352 up-regulated and 174 down-regulated circRNAs at least 2-fold change compared to control(264). HOTAIR expression is down-regulated in HPV-16 positive cervical cancer samples compared to HPV-negative normal samples(265). In cervical cancer, HOTAIR directly interacts with HPV-16 E7 and is inactivated by the HPV protein.

MALAT1 has been the most extensively studied HPV-regulated lncRNA. MALAT1 expression is increased in HPV-16 and HPV-18 positive human cervical cancer samples compared to normal human cervical samples(141, 218). Additionally, separating HPV-positive and HPV-negative samples, increased

MALAT1 expression correlated with HPV-positive status. In cervical cancer cells, MALAT1 has been shown to be involved in the regulation of miRNAs to contribute to carcinogenesis(218, 266). One group specifically showed that in high-risk HPV cervical cancer, MALAT1 expression is higher in radioresistant than radiosensitive cancer cases and that MALAT1 sponges miR-145 to mediate this radioresistance(267). Decreasing MALAT1 expression in cervical cancer cells altered proliferation and invasion(216). Knockdown of HPV-16 E6/E7 led to decreased MALAT1 expression, showing a potential E6/E7 regulation of MALAT1(141). Additionally, MALAT1 is up-regulated in oral keratinocytes expressing HPV-16 E6/E7(268).

Overall, host lncRNAs are altered by individual expression of each of the HPV oncoproteins, co-expression of both oncoproteins, as well as expression of the entire HPV genome. As miRNAs are altered by HPV to regulate the viral life cycle and additionally cause alterations that contribute to cancer, it is speculated that many lncRNAs are being altered by HPV for similar functions.

B. HPV regulation of lncRNAs in cancer - Clinical implications

A clinical trial sponsored in China is currently recruiting uterine cervical adenocarcinoma patients to analyze multi-omics results of patients with and without HPV infection, with the goal of revealing the role of HPV integration in molecular mechanism(s) of tumorigenesis and prognosis of patients. They will analyze for changes in transcriptomics (lncRNA, miRNA, mRNA), whole-exome (DNA), and metabolomics (NCT03742869).

VII. Oncogenic lncRNA FAM83H-AS1 (also known as 'onco-lncRNA-3')

In 2014, LCAL7 (lung cancer associated long non-coding RNA 7), which contains at least part of FAM83H-AS1, ENST00000533004, ENST00000534398, and ENST00000435097 in the genome, was first identified as up-regulated in lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) human tissues compared to controls. LCAL7 is located in chr8: 144,823,553-144,828,510 and a partial sequence of LCAL7 was identified by 5' and 3' RACE in a lung cancer cell line and can be found in GenBank (accession number: KF773847)(269). Overall, LCAL7 (containing a portion of FAM83H-AS1) is up-regulated in lung cancer.

In 2015, 'onco-lncRNA-3' expression was further analyzed using publicly available TCGA RNA-seq data from tumors and matched normal tissues. Onco-lncRNA-3 is up-regulated in breast invasive

carcinoma (BRCA), colorectal and rectal cancer (CRC), lung adenocarcinoma (LUAD), and lung squamous cell carcinoma (LUSC). Additionally, analysis showed 'outlier' expression in head and neck squamous cell carcinoma (HNSC); 'outlier' referred to subset of patients from the certain cancer type exhibiting consistent gene expression changes. Onco-lncRNA-3 was not found to be altered in bladder urothelial carcinoma (BLCA), kidney renal cell carcinoma (KIRC), or uterine corpus endometrial carcinoma (UCEC). Functionally, onco-lncRNA-3 exhibited a 'guilt by association' effect as it was co-expressed with many protein coding genes involved in cell cycle regulation, such that it was speculated that onco-lncRNA-3 plays a role in cell cycle progression. Onco-lncRNA-3 was shown to play a role in S-phase cell cycle in cancer as knockdown of the lncRNA in color cancer and lung cancer cells resulted in decreased EdU incorporation compared to control(133). In this study, onco-lncRNA-3 was found to be up-regulated in breast cancer, colorectal cancer, and lung cancer and plays a role in cell cycle progression.

Yang et al. identified FAM83H-AS1 (family with sequence similarity 83 member H-antisense 1) as the most up-regulated lncRNA in their RNA-seq analysis of breast cancer tissues compared to normal tissues. Additionally, analysis of 626 breast cancer samples compared to 105 non-cancerous breast tissues showed a significantly higher FAM83H-AS1 RPKM value in the cancerous tissues. Analyzing TCGA data from luminal subtype breast cancer patients revealed higher expression of FAM83H-AS1 correlates with poorer overall survival. Analyzing a TCGA cohort of breast cancer patients (containing all subtypes) revealed high expression of FAM83H-AS1 correlated with poorer overall survival(270). Taken together, FAM83H-AS1 is an independent prognostic marker in breast cancer.

Analysis of multiple colorectal cancer microarrays for differential expression of lncRNAs showed FAM83H-AS1 was the top up-regulated lncRNA analyzed in colorectal cancer (CRC) tissues compared to adjacent normal tissues. Further analysis of human colorectal cancer cell lines confirmed expression of FAM83H-AS1 in all but one cell line analyzed. High FAM83H-AS1 expression in CRC patients correlated with poorer overall survival than those with low FAM83H-AS1 expression(271). FAM83H-AS1 could be a potential prognosis biomarker in colorectal cancer patients.

RNA-seq analysis of LUAD tissues compared to controls in TCGA showed increased FAM83H-AS1. Further studies analyzing FAM83H-AS1 expression in tumor versus adjacent normal tissue of the same lineage revealed significant up-regulation of FAM83H-AS1 in LUAD, lung large cell carcinoma (LULC), LUSC, bladder cancer, breast cancer, gastric cancer, head and neck cancer, and prostate cancer while no significant differences in kidney cancer, liver cancer, or thyroid cancer was observed in their analysis(272). Notably, data from a different study showed gastric cancer patients in TCGA revealed no

consistent differential expression of FAM83H-AS1(273). In lung cancer clinical patients, high expression of FAM83H-AS1 correlated with poorer patient survival. Functionally, knockdown of FAM83H-AS1 expression in lung cancer cell lines decreased proliferation, migration, and invasion as well as decreased percentage of cells in S-phase of the cell cycle. Notably, proliferation was affected in most EGFR-mutated lung cancer cells. FAM83H-AS1 is primarily localized in the nucleus of lung cancer cells. Mechanistically, knockdown of FAM83H-AS1 decreased protein expression of MET, EGFR and their downstream proteins AKT and ERK1/2 in lung cancer cell lines. Notably, knockdown of FAM83H-AS1 regulated MET at the transcriptional and translational level(272). In lung cancer, FAM83H-AS1 could potentially be used as a diagnostic and prognostic marker or a therapeutic target to impact MET/EGFR signaling involved with cell proliferation and invasion.

RNA-seq data in TCGA of pancreas-adenocarcinoma ductal type (PDA) was analyzed for differential lncRNA expression. Analyzing relative enrichment in the stroma or epithelium identified FAM83H-AS1 as one of the top epithelial lncRNAs altered in their study. FAM83H-AS1 and the nearby protein coding gene FAM83H likely share the same promoter region, but are transcribed in opposite directions. FAM83H-AS1 is located in a genomic region frequently amplified in PDA (8q23.3-8q24.3), and its expression does in fact correlate significantly with amplification. Analysis indicates that it has four exons, and is located in an actively transcribed region in a pancreatic cancer cell line. FAM83H-AS1 expression is shown to be up-regulated in PDA lines as well as a breast cancer cell line. Clinically, high expression of FAM83H-AS1 correlates with poorer overall survival in PDA. Knockdown of FAM83H-AS1 in a pancreatic cancer cell line by two different siRNAs revealed 719 genes differentially expressed in cells with siRNA1 and siRNA2. Regulatory network analysis predicted 146 target genes of FAM83H-AS1 (78 positive/activated and 68 negative/inhibited). Of these target genes, the negatively regulated targets were more associated with benign processes while positively regulated targets are associated with more malignant processes(274). Thus, FAM83H-AS1 is involved with malignant progression in pancreatic cancer.

Bi et al. identified FAM83H-AS1 is up-regulated in glioma tissues and cell lines compared to controls. FAM83H-AS1 expression significantly correlates with tumor grade. Patients with high FAM83H-AS1 expression were associated with worse overall survival in glioma. High-grade tumors FAM83H-AS1 expression correlates with patients' prognosis. FAM83H-AS1 expression could be an independent prognostic factor in glioma. Functionally, for the first time, a group expressed FAM83H-AS1 to observe functional alterations in the cells. Over-expression of FAM83H-AS1 in glioma cells increased cell viability and colony formation. Knockdown of FAM83H-AS1 in glioma cells reduced cell viability, colony

formation, percentage of cells in S-phase of the cell cycle, and cyclin-dependent kinases (CDK2, CDK4, CDK6), while it increased cellular apoptosis in glioma cells. Additionally, knockdown of FAM83H-AS1 in glioma cells resulted in increased CDK inhibitors CDKN1A (also known as p21) mRNA expression. FAM83H-AS1 is primarily localized in the nucleus and contributes to epigenetic transcriptional regulation of p21 expression by binding to EZH2 in glioma cells. Decrease in p21 expression potentially contributes to FAM83H-AS1 oncogenic function(275). Together, FAM83H-AS1 could be a prognostic marker in glioma patients, and/or therapeutically targeting FAM83H-AS1 and in turn modulating its downstream targets might improve glioma patient outcome.

FAM83H-AS1 expression is up-regulated in colorectal cancer (CRC) tissues and cells. In CRC clinical patients, FAM83H-AS1 expression correlates with tumor stage and size and high expression of FAM83H-AS1 is associated with decreased overall survival. Functionally, knockdown of FAM83H-AS1 in colorectal cancer cells with specific short-hairpin RNA (shRNA) led to decreased proliferation, colony formation, and migration as well as increased cell apoptosis. Additionally, knockdown of FAM83H-AS1 led to a decrease in Notch1 mRNA expression. Treatment of colorectal cancer cells with a widely used gamma-secretase inhibitor (DAPT) that represses receptors and ligands in the Notch signaling pathway (including Notch1) decreased FAM83H-AS1 expression. These results are indicative that FAM83H-AS1 regulates through the Notch pathway.(276) Overall, analysis of FAM83H-AS1 expression in colorectal cancer tissues can act as an independent prognostic factor in colorectal carcinoma. Therapeutically targeting FAM83H-AS1 to modulate downstream Notch signaling might be clinically useful as Notch signaling is involved with regulation of cell proliferation, differentiation, and death.

Our group identified FAM83H-AS1 expression is up-regulated in cervical cancer tissues and cells, and clinical patients with high FAM83H-AS1 expression exhibited a significantly worse overall survival than those with low FAM83H-AS1 expression. Notably, in both cervical and HNSCC cells, FAM83H-AS1 expression was significantly higher in HPV-positive compared to HPV-negative samples. Mechanistically, FAM83H-AS1 expression is regulated in a p53-independent and at least partially by HPV-16 E6 expression and p300 in cervical cancer. In cervical cancer cells, FAM83H-AS1 expression is primarily localized in the nucleus and does not appear to regulate its nearby protein coding gene FAM83H. Knockdown of FAM83H-AS1 expression in cervical cancer cells decreased proliferation (and increased percentage of cells in the S-phase of cell cycle), migration as well as increased apoptosis (Chapter 2). Together, FAM83H-AS1 could be a potential prognostic marker in cervical cancer, further studies could reveal that FAM83H-AS1 could be a diagnostic marker for HPV infection progression to cervical cancer, and therapeutically targeting FAM83H-AS1 could improve patient outcome.

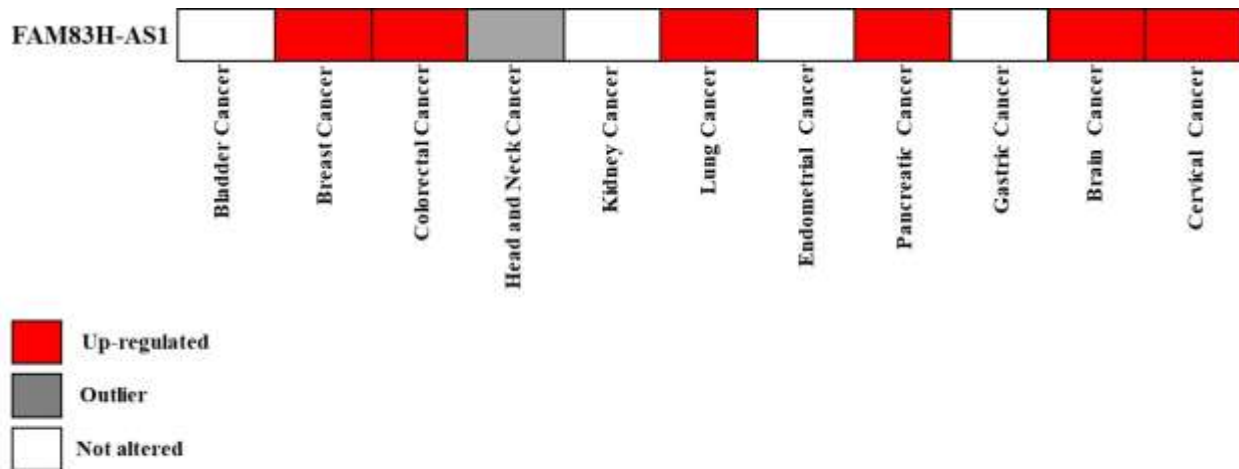


Figure 13: Heat map of FAM83H-AS1 expression alterations in multiple cancer types. Adapted from: *Cabanski et al., RNA Biology, 2015.*

Overall, FAM83H-AS1 is up-regulated in many lung(133, 269, 272), breast(133, 270, 272), colorectal(133, 271, 276), pancreatic(274), glioma brain(275), bladder(272), prostate(272), and cervical cancers (Figure 13 and Chapter 2). FAM83H-AS1 expression is upregulated in certain gastric cancer(272) and HNSCC(272) studies, while others show variable alterations in gastric(273) and HNSCC(133) tissues. Thus, FAM83H-AS1 could potentially be a diagnostic marker in those cancers. High expression of FAM83H-AS1 correlates with worse overall survival in breast(270), colorectal(271, 276), lung(272), pancreatic(274), glioma brain(275), and cervical cancers (Chapter 2); therefore, FAM83H-AS1 could potentially be a prognostic marker in those cancer types. FAM83H-AS1 is mainly localized in the epithelium in pancreatic cancer cells(274) and specifically in the nucleus in lung cancer(272), cervical cancer (Chapter 2), and glioma brain cancer(275) cells. FAM83H-AS1 expression is up-regulated at least in part by HPV-16 E6 and p300 in cervical cancer cells (Chapter 2). Downstream, FAM83H-AS1 regulates Notch signaling in colorectal cancer(276), MET/EGFR signaling in lung cancer(272), and p21 expression in glioma(275). Further studies need to be conducted to determine if these upstream and downstream targets involved with FAM83H-AS1 expression are consistent between cancer types, and novel upstream and downstream regulators of FAM83H-AS1 expression need to be elucidated to increase therapeutic potential of targeting FAM83H-AS1.

References:

1. **Jacob F, Monod J.** 1961. Genetic regulatory mechanisms in the synthesis of proteins. *J Mol Biol* **3**:318-356.
2. **Fox GE.** 2010. Origin and evolution of the ribosome. *Cold Spring Harb Perspect Biol* **2**:a003483.
3. **Kruger K, Grabowski PJ, Zaug AJ, Sands J, Gottschling DE, Cech TR.** 1982. Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. *Cell* **31**:147-157.
4. **Garriga G, Lambowitz AM.** 1984. RNA splicing in neurospora mitochondria: self-splicing of a mitochondrial intron in vitro. *Cell* **39**:631-641.
5. **Anonymous.** 2012. An integrated encyclopedia of DNA elements in the human genome. *Nature* **489**:57-74.
6. **Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F, Parker SC, Sabo PJ, Sandstrom R, Shafer A, Vetrie D, et al.** 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**:799-816.
7. **Bazzini AA, Johnstone TG, Christiano R, Mackowiak SD, Obermayer B, Fleming ES, Vejnar CE, Lee MT, Rajewsky N, Walther TC, Giraldez AJ.** 2014. Identification of small ORFs in vertebrates using ribosome footprinting and evolutionary conservation. *Embo j* **33**:981-993.
8. **Galindo MI, Pueyo JI, Fouix S, Bishop SA, Couso JP.** 2007. Peptides encoded by short ORFs control development and define a new eukaryotic gene family. *PLoS Biol* **5**:e106.
9. **Lee Y, Rio DC.** 2015. Mechanisms and Regulation of Alternative Pre-mRNA Splicing. *Annu Rev Biochem* **84**:291-323.
10. **Napoli C, Lemieux C, Jorgensen R.** 1990. Introduction of a Chimeric Chalcone Synthase Gene into Petunia Results in Reversible Co-Suppression of Homologous Genes in trans. *Plant Cell* **2**:279-289.
11. **Fire A, Xu S, Montgomery MK, Kostas SA, Driver SE, Mello CC.** 1998. Potent and specific genetic interference by double-stranded RNA in *Caenorhabditis elegans*. *Nature* **391**:806-811.
12. **Iwasaki YW, Siomi MC, Siomi H.** 2015. PIWI-Interacting RNA: Its Biogenesis and Functions. *Annu Rev Biochem* **84**:405-433.
13. **Wen YZ, Zheng LL, Liao JY, Wang MH, Wei Y, Guo XM, Qu LH, Ayala FJ, Lun ZR.** 2011. Pseudogene-derived small interference RNAs regulate gene expression in African Trypanosoma brucei. *Proc Natl Acad Sci U S A* **108**:8345-8350.
14. **Guo X, Lin M, Rockowitz S, Lachman HM, Zheng D.** 2014. Characterization of human pseudogene-derived non-coding RNAs for functional potential. *PLoS One* **9**:e93972.
15. **Okazaki Y, Furuno M, Kasukawa T, Adachi J, Bono H, Kondo S, Nikaido I, Osato N, Saito R, Suzuki H, Yamanaka I, Kiyosawa H, Yagi K, Tomaru Y, Hasegawa Y, Nogami A, Schonbach C, Gojobori T, Baldarelli R, Hill DP, Bult C, Hume DA, Quackenbush J, Schriml LM, Kanapin A, Matsuda H, Batalov S, Beisel KW, Blake JA, Bradt D, Brusica V, Chothia C, Corbani LE, Cousins S, Dalla E, Dragani TA, Fletcher CF, Forrest A, Frazer KS, Gaasterland T, Gariboldi M, Gissi C, Godzik A, Gough J, Grimmond S, Gustincich S, Hirokawa N, Jackson IJ, Jarvis ED, Kanai A, et al.** 2002. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs. *Nature* **420**:563-573.
16. **Lee RC, Feinbaum RL, Ambros V.** 1993. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**:843-854.

17. **Calin GA, Dumitru CD, Shimizu M, Bichi R, Zupo S, Noch E, Aldler H, Rattan S, Keating M, Rai K, Rassenti L, Kipps T, Negrini M, Bullrich F, Croce CM.** 2002. Frequent deletions and down-regulation of micro- RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. *Proc Natl Acad Sci U S A* **99**:15524-15529.
18. **Berezikov E, Guryev V, van de Belt J, Wienholds E, Plasterk RH, Cuppen E.** 2005. Phylogenetic shadowing and computational identification of human microRNA genes. *Cell* **120**:21-24.
19. **Lau NC, Lim LP, Weinstein EG, Bartel DP.** 2001. An abundant class of tiny RNAs with probable regulatory roles in *Caenorhabditis elegans*. *Science* **294**:858-862.
20. **Lagos-Quintana M, Rauhut R, Lendeckel W, Tuschl T.** 2001. Identification of novel genes coding for small expressed RNAs. *Science* **294**:853-858.
21. **Lee RC, Ambros V.** 2001. An extensive class of small RNAs in *Caenorhabditis elegans*. *Science* **294**:862-864.
22. **Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB.** 2003. Prediction of mammalian microRNA targets. *Cell* **115**:787-798.
23. **Kozomara A, Griffiths-Jones S.** 2014. miRBase: annotating high confidence microRNAs using deep sequencing data. *Nucleic Acids Res* **42**:D68-73.
24. **Hamilton AJ, Baulcombe DC.** 1999. A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* **286**:950-952.
25. **Liu J, Valencia-Sanchez MA, Hannon GJ, Parker R.** 2005. MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies. *Nat Cell Biol* **7**:719-723.
26. **Lee Y, Kim M, Han J, Yeom KH, Lee S, Baek SH, Kim VN.** 2004. MicroRNA genes are transcribed by RNA polymerase II. *Embo j* **23**:4051-4060.
27. **Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, Kim VN.** 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**:415-419.
28. **Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN.** 2004. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* **18**:3016-3027.
29. **Basyuk E, Suavet F, Doglio A, Bordonne R, Bertrand E.** 2003. Human let-7 stem-loop precursors harbor features of RNase III cleavage products. *Nucleic Acids Res* **31**:6593-6597.
30. **Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U.** 2004. Nuclear export of microRNA precursors. *Science* **303**:95-98.
31. **Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ.** 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**:231-235.
32. **Yi R, Qin Y, Macara IG, Cullen BR.** 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**:3011-3016.
33. **Bohnsack MT, Czaplinski K, Gorlich D.** 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *Rna* **10**:185-191.
34. **Bernstein E, Caudy AA, Hammond SM, Hannon GJ.** 2001. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**:363-366.
35. **Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, Baillie DL, Fire A, Ruvkun G, Mello CC.** 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**:23-34.
36. **Hutvagner G, McLachlan J, Pasquinelli AE, Balint E, Tuschl T, Zamore PD.** 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**:834-838.
37. **Ketting RF, Fischer SE, Bernstein E, Sijen T, Hannon GJ, Plasterk RH.** 2001. Dicer functions in RNA interference and in synthesis of small RNA involved in developmental timing in *C. elegans*. *Genes Dev* **15**:2654-2659.
38. **Knight SW, Bass BL.** 2001. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293**:2269-2271.

39. **O'Toole AS, Miller S, Haines N, Zink MC, Serra MJ.** 2006. Comprehensive thermodynamic analysis of 3' double-nucleotide overhangs neighboring Watson-Crick terminal base pairs. *Nucleic Acids Res* **34**:3338-3344.
40. **Khvorova A, Reynolds A, Jayasena SD.** 2003. Functional siRNAs and miRNAs exhibit strand bias. *Cell* **115**:209-216.
41. **Schwarz DS, Hutvagner G, Du T, Xu Z, Aronin N, Zamore PD.** 2003. Asymmetry in the assembly of the RNAi enzyme complex. *Cell* **115**:199-208.
42. **Huang CJ, Nguyen PN, Choo KB, Sugii S, Wee K, Cheong SK, Kamarul T.** 2014. Frequent co-expression of miRNA-5p and -3p species and cross-targeting in induced pluripotent stem cells. *Int J Med Sci* **11**:824-833.
43. **Bartel DP.** 2004. MicroRNAs: genomics, biogenesis, mechanism, and function. *Cell* **116**:281-297.
44. **Yekta S, Shih IH, Bartel DP.** 2004. MicroRNA-directed cleavage of HOXB8 mRNA. *Science* **304**:594-596.
45. **Eulalio A, Huntzinger E, Nishihara T, Rehwinkel J, Fauser M, Izaurralde E.** 2009. Deadenylation is a widespread effect of miRNA regulation. *Rna* **15**:21-32.
46. **Kim YK, Kim B, Kim VN.** 2016. Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. *Proc Natl Acad Sci U S A* **113**:E1881-1889.
47. **Okamura K, Hagen JW, Duan H, Tyler DM, Lai EC.** 2007. The mirtron pathway generates microRNA-class regulatory RNAs in *Drosophila*. *Cell* **130**:89-100.
48. **Ruby JG, Jan CH, Bartel DP.** 2007. Intronic microRNA precursors that bypass Drosha processing. *Nature* **448**:83-86.
49. **Berezikov E, Chung WJ, Willis J, Cuppen E, Lai EC.** 2007. Mammalian mirtron genes. *Mol Cell* **28**:328-336.
50. **Babiarz JE, Ruby JG, Wang Y, Bartel DP, Blelloch R.** 2008. Mouse ES cells express endogenous shRNAs, siRNAs, and other Microprocessor-independent, Dicer-dependent small RNAs. *Genes Dev* **22**:2773-2785.
51. **Glazov EA, Cottee PA, Barris WC, Moore RJ, Dalrymple BP, Tizard ML.** 2008. A microRNA catalog of the developing chicken embryo identified by a deep sequencing approach. *Genome Res* **18**:957-964.
52. **Ladewig E, Okamura K, Flynt AS, Westholm JO, Lai EC.** 2012. Discovery of hundreds of mirtrons in mouse and human small RNA data. *Genome Res* **22**:1634-1645.
53. **Flynt AS, Greimann JC, Chung WJ, Lima CD, Lai EC.** 2010. MicroRNA biogenesis via splicing and exosome-mediated trimming in *Drosophila*. *Mol Cell* **38**:900-907.
54. **Havens MA, Reich AA, Duelli DM, Hastings ML.** 2012. Biogenesis of mammalian microRNAs by a non-canonical processing pathway. *Nucleic Acids Res* **40**:4626-4640.
55. **Ender C, Krek A, Friedlander MR, Beitzinger M, Weinmann L, Chen W, Pfeffer S, Rajewsky N, Meister G.** 2008. A human snoRNA with microRNA-like functions. *Mol Cell* **32**:519-528.
56. **Brameier M, Herwig A, Reinhardt R, Walter L, Gruber J.** 2011. Human box C/D snoRNAs with miRNA like functions: expanding the range of regulatory RNAs. *Nucleic Acids Res* **39**:675-686.
57. **Scott MS, Avolio F, Ono M, Lamond AI, Barton GJ.** 2009. Human miRNA precursors with box H/ACA snoRNA features. *PLoS Comput Biol* **5**:e1000507.
58. **Scott MS, Ono M.** 2011. From snoRNA to miRNA: Dual function regulatory non-coding RNAs. *Biochimie* **93**:1987-1992.
59. **Ono M, Scott MS, Yamada K, Avolio F, Barton GJ, Lamond AI.** 2011. Identification of human miRNA precursors that resemble box C/D snoRNAs. *Nucleic Acids Res* **39**:3879-3891.
60. **Cole C, Sobala A, Lu C, Thatcher SR, Bowman A, Brown JW, Green PJ, Barton GJ, Hutvagner G.** 2009. Filtering of deep sequencing data reveals the existence of abundant Dicer-dependent small RNAs derived from tRNAs. *Rna* **15**:2147-2160.
61. **Haussecker D, Huang Y, Lau A, Parameswaran P, Fire AZ, Kay MA.** 2010. Human tRNA-derived small RNAs in the global regulation of RNA silencing. *Rna* **16**:673-695.

62. **Bowden RJ, Simas JP, Davis AJ, Efstathiou S.** 1997. Murine gammaherpesvirus 68 encodes tRNA-like sequences which are expressed during latency. *J Gen Virol* **78** (Pt 7):1675-1687.
63. **Reese TA, Xia J, Johnson LS, Zhou X, Zhang W, Virgin HW.** 2010. Identification of novel microRNA-like molecules generated from herpesvirus and host tRNA transcripts. *J Virol* **84**:10344-10353.
64. **Bogerd HP, Karnowski HW, Cai X, Shin J, Pohlers M, Cullen BR.** 2010. A mammalian herpesvirus uses noncanonical expression and processing mechanisms to generate viral MicroRNAs. *Mol Cell* **37**:135-142.
65. **Shapiro JS, Varble A, Pham AM, Tenover BR.** 2010. Noncanonical cytoplasmic processing of viral microRNAs. *Rna* **16**:2068-2074.
66. **Shapiro JS, Langlois RA, Pham AM, Tenover BR.** 2012. Evidence for a cytoplasmic microprocessor of pri-miRNAs. *Rna* **18**:1338-1346.
67. **Cazalla D, Xie M, Steitz JA.** 2011. A primate herpesvirus uses the integrator complex to generate viral microRNAs. *Mol Cell* **43**:982-992.
68. **Chirayil R, Kincaid RP, Dahlke C, Kuny CV, Dalken N, Spohn M, Lawson B, Grundhoff A, Sullivan CS.** 2018. Identification of virus-encoded microRNAs in divergent Papillomaviruses. *PLoS Pathog* **14**:e1007156.
69. **Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ.** 2010. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**:584-589.
70. **Cifuentes D, Xue H, Taylor DW, Patnode H, Mishima Y, Cheloufi S, Ma E, Mane S, Hannon GJ, Lawson ND, Wolfe SA, Giraldez AJ.** 2010. A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **328**:1694-1698.
71. **Yang JS, Lai EC.** 2010. Dicer-independent, Ago2-mediated microRNA biogenesis in vertebrates. *Cell Cycle* **9**:4455-4460.
72. **Yang JS, Maurin T, Robine N, Rasmussen KD, Jeffrey KL, Chandwani R, Papapetrou EP, Sadelain M, O'Carroll D, Lai EC.** 2010. Conserved vertebrate mir-451 provides a platform for Dicer-independent, Ago2-mediated microRNA biogenesis. *Proc Natl Acad Sci U S A* **107**:15163-15168.
73. **McCloskey A, Taniguchi I, Shinmyozu K, Ohno M.** 2012. hnRNP C tetramer measures RNA length to classify RNA polymerase II transcripts for export. *Science* **335**:1643-1646.
74. **Xie M, Li M, Vilborg A, Lee N, Shu MD, Yartseva V, Sestan N, Steitz JA.** 2013. Mammalian 5'-capped microRNA precursors that generate a single microRNA. *Cell* **155**:1568-1580.
75. **Garzon R, Calin GA, Croce CM.** 2009. MicroRNAs in Cancer. *Annu Rev Med* **60**:167-179.
76. **Lee YS, Dutta A.** 2009. MicroRNAs in cancer. *Annu Rev Pathol* **4**:199-227.
77. **Di Leva G, Garofalo M, Croce CM.** 2014. MicroRNAs in cancer. *Annu Rev Pathol* **9**:287-314.
78. **Peng Y, Croce CM.** 2016. The role of MicroRNAs in human cancer. *Signal Transduct Target Ther* **1**:15004.
79. **Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, Peck D, Sweet-Cordero A, Ebert BL, Mak RH, Ferrando AA, Downing JR, Jacks T, Horvitz HR, Golub TR.** 2005. MicroRNA expression profiles classify human cancers. *Nature* **435**:834-838.
80. **Gaur A, Jewell DA, Liang Y, Ridzon D, Moore JH, Chen C, Ambros VR, Israel MA.** 2007. Characterization of microRNA expression levels and their biological correlates in human cancer cell lines. *Cancer Res* **67**:2456-2468.
81. **Kumar MS, Lu J, Mercer KL, Golub TR, Jacks T.** 2007. Impaired microRNA processing enhances cellular transformation and tumorigenesis. *Nat Genet* **39**:673-677.
82. **Arvey A, Larsson E, Sander C, Leslie CS, Marks DS.** 2010. Target mRNA abundance dilutes microRNA and siRNA activity. *Mol Syst Biol* **6**:363.
83. **Calin GA, Sevignani C, Dumitru CD, Hyslop T, Noch E, Yendamuri S, Shimizu M, Rattan S, Bullrich F, Negrini M, Croce CM.** 2004. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. *Proc Natl Acad Sci U S A* **101**:2999-3004.

84. **Lehmann U, Hasemeier B, Christgen M, Muller M, Romermann D, Langer F, Kreipe H.** 2008. Epigenetic inactivation of microRNA gene hsa-mir-9-1 in human breast cancer. *J Pathol* **214**:17-24.
85. **Lujambio A, Ropero S, Ballestar E, Fraga MF, Cerrato C, Setien F, Casado S, Suarez-Gauthier A, Sanchez-Cespedes M, Git A, Spiteri I, Das PP, Caldas C, Miska E, Esteller M.** 2007. Genetic unmasking of an epigenetically silenced microRNA in human cancer cells. *Cancer Res* **67**:1424-1429.
86. **Fabbri M, Garzon R, Cimmino A, Liu Z, Zanesi N, Callegari E, Liu S, Alder H, Costinean S, Fernandez-Cymering C, Volinia S, Guler G, Morrison CD, Chan KK, Marcucci G, Calin GA, Huebner K, Croce CM.** 2007. MicroRNA-29 family reverts aberrant methylation in lung cancer by targeting DNA methyltransferases 3A and 3B. *Proc Natl Acad Sci U S A* **104**:15805-15810.
87. **Dews M, Homayouni A, Yu D, Murphy D, Seignani C, Wentzel E, Furth EE, Lee WM, Enders GH, Mendell JT, Thomas-Tikhonenko A.** 2006. Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster. *Nat Genet* **38**:1060-1065.
88. **He L, He X, Lim LP, de Stanchina E, Xuan Z, Liang Y, Xue W, Zender L, Magnus J, Ridzon D, Jackson AL, Linsley PS, Chen C, Lowe SW, Cleary MA, Hannon GJ.** 2007. A microRNA component of the p53 tumour suppressor network. *Nature* **447**:1130-1134.
89. **He H, Jazdzewski K, Li W, Liyanarachchi S, Nagy R, Volinia S, Calin GA, Liu CG, Franssila K, Suster S, Kloos RT, Croce CM, de la Chapelle A.** 2005. The role of microRNA genes in papillary thyroid carcinoma. *Proc Natl Acad Sci U S A* **102**:19075-19080.
90. **Fontana L, Fiori ME, Albin S, Cifaldi L, Giovinnazzi S, Forloni M, Boldrini R, Donfrancesco A, Federici V, Giacomini P, Peschle C, Fruci D.** 2008. Antagomir-17-5p abolishes the growth of therapy-resistant neuroblastoma through p21 and BIM. *PLoS One* **3**:e2236.
91. **Buscaglia LE, Li Y.** 2011. Apoptosis and the target genes of microRNA-21. *Chin J Cancer* **30**:371-380.
92. **Loffler D, Brocke-Heidrich K, Pfeifer G, Stocsits C, Hackermuller J, Kretzschmar AK, Burger R, Gramatzki M, Blumert C, Bauer K, Cvijic H, Ullmann AK, Stadler PF, Horn F.** 2007. Interleukin-6 dependent survival of multiple myeloma cells involves the Stat3-mediated induction of microRNA-21 through a highly conserved enhancer. *Blood* **110**:1330-1333.
93. **Mattiske S, Suetani RJ, Neilsen PM, Callen DF.** 2012. The oncogenic role of miR-155 in breast cancer. *Cancer Epidemiol Biomarkers Prev* **21**:1236-1243.
94. **Bayraktar R, Van Roosbroeck K.** 2018. miR-155 in cancer drug resistance and as target for miRNA-based therapeutics. *Cancer Metastasis Rev* **37**:33-44.
95. **Wang X, Cao L, Wang Y, Wang X, Liu N, You Y.** 2012. Regulation of let-7 and its target oncogenes (Review). *Oncol Lett* **3**:955-960.
96. **Slabakova E, Culig Z, Remsik J, Soucek K.** 2017. Alternative mechanisms of miR-34a regulation in cancer. *Cell Death Dis* **8**:e3100.
97. **Lin Z, Flemington EK.** 2011. miRNAs in the pathogenesis of oncogenic human viruses. *Cancer Lett* **305**:186-199.
98. **Martinez I, Gardiner AS, Board KF, Monzon FA, Edwards RP, Khan SA.** 2008. Human papillomavirus type 16 reduces the expression of microRNA-218 in cervical carcinoma cells. *Oncogene* **27**:2575-2582.
99. **Wang X, Wang HK, McCoy JP, Banerjee NS, Rader JS, Broker TR, Meyers C, Chow LT, Zheng ZM.** 2009. Oncogenic HPV infection interrupts the expression of tumor-suppressive miR-34a through viral oncoprotein E6. *Rna* **15**:637-647.
100. **Samols MA, Skalsky RL, Maldonado AM, Riva A, Lopez MC, Baker HV, Renne R.** 2007. Identification of cellular genes targeted by KSHV-encoded microRNAs. *PLoS Pathog* **3**:e65.
101. **Pfeffer S, Zavolan M, Grasser FA, Chien M, Russo JJ, Ju J, John B, Enright AJ, Marks D, Sander C, Tuschl T.** 2004. Identification of virus-encoded microRNAs. *Science* **304**:734-736.

102. **Taraboletti G, Benelli R, Borsotti P, Rusnati M, Presta M, Giavazzi R, Ruco L, Albini A.** 1999. Thrombospondin-1 inhibits Kaposi's sarcoma (KS) cell and HIV-1 Tat-induced angiogenesis and is poorly expressed in KS lesions. *J Pathol* **188**:76-81.
103. **Guttman M, Amit I, Garber M, French C, Lin MF, Feldser D, Huarte M, Zuk O, Carey BW, Cassady JP, Cabili MN, Jaenisch R, Mikkelsen TS, Jacks T, Hacohen N, Bernstein BE, Kellis M, Regev A, Rinn JL, Lander ES.** 2009. Chromatin signature reveals over a thousand highly conserved large non-coding RNAs in mammals. *Nature* **458**:223-227.
104. **Miao Y, Xu SY, Chen LS, Liang GY, Pu YP, Yin LH.** 2017. Trends of long noncoding RNA research from 2007 to 2016: a bibliometric analysis. *Oncotarget* **8**:83114-83127.
105. **Kashi K, Henderson L, Bonetti A, Carninci P.** 2016. Discovery and functional analysis of lncRNAs: Methodologies to investigate an uncharacterized transcriptome. *Biochim Biophys Acta* **1859**:3-15.
106. **Quinn JJ, Chang HY.** 2016. Unique features of long non-coding RNA biogenesis and function. *Nat Rev Genet* **17**:47-62.
107. **Wu H, Yang L, Chen LL.** 2017. The Diversity of Long Noncoding RNAs and Their Generation. *Trends Genet* **33**:540-552.
108. **Sun Q, Hao Q, Prasanth KV.** 2018. Nuclear Long Noncoding RNAs: Key Regulators of Gene Expression. *Trends Genet* **34**:142-157.
109. **Grzechnik P, Tan-Wong SM, Proudfoot NJ.** 2014. Terminate and make a loop: regulation of transcriptional directionality. *Trends Biochem Sci* **39**:319-327.
110. **Salzman J, Gawad C, Wang PL, Lacayo N, Brown PO.** 2012. Circular RNAs are the predominant transcript isoform from hundreds of human genes in diverse cell types. *PLoS One* **7**:e30733.
111. **Rinn JL, Chang HY.** 2012. Genome regulation by long noncoding RNAs. *Annu Rev Biochem* **81**:145-166.
112. **Herrera-Solorio AM, Armas-Lopez L, Arrieta O, Zuniga J, Pina-Sanchez P, Avila-Moreno F.** 2017. Histone code and long non-coding RNAs (lncRNAs) aberrations in lung cancer: implications in the therapy response. *Clin Epigenetics* **9**:98.
113. **Ning Q, Li Y, Wang Z, Zhou S, Sun H, Yu G.** 2017. The Evolution and Expression Pattern of Human Overlapping lncRNA and Protein-coding Gene Pairs. *Sci Rep* **7**:42775.
114. **Balbin OA, Malik R, Dhanasekaran SM, Prensner JR, Cao X, Wu YM, Robinson D, Wang R, Chen G, Beer DG, Nesvizhskii AI, Chinnaiyan AM.** 2015. The landscape of antisense gene expression in human cancers. *Genome Res* **25**:1068-1079.
115. **Carrieri C, Cimatti L, Biagioli M, Beugnet A, Zucchelli S, Fedele S, Pesce E, Ferrer I, Collavin L, Santoro C, Forrest AR, Carninci P, Biffo S, Stupka E, Gustincich S.** 2012. Long non-coding antisense RNA controls Uchl1 translation through an embedded SINEB2 repeat. *Nature* **491**:454-457.
116. **Kiyosawa H, Mise N, Iwase S, Hayashizaki Y, Abe K.** 2005. Disclosing hidden transcripts: mouse natural sense-antisense transcripts tend to be poly(A) negative and nuclear localized. *Genome Res* **15**:463-474.
117. **Morrissy AS, Griffith M, Marra MA.** 2011. Extensive relationship between antisense transcription and alternative splicing in the human genome. *Genome Res* **21**:1203-1212.
118. **Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES, Rinn JL.** 2009. Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. *Proc Natl Acad Sci U S A* **106**:11667-11672.
119. **Marques AC, Hughes J, Graham B, Kowalczyk MS, Higgs DR, Ponting CP.** 2013. Chromatin signatures at transcriptional start sites separate two equally populated yet distinct classes of intergenic long noncoding RNAs. *Genome Biol* **14**:R131.
120. **Ounzain S, Pezzuto I, Micheletti R, Burdet F, Sheta R, Nemir M, Gonzales C, Sarre A, Alexanian M, Blow MJ, May D, Johnson R, Dauvillier J, Pennacchio LA, Pedrazzini T.** 2014.

- Functional importance of cardiac enhancer-associated noncoding RNAs in heart development and disease. *J Mol Cell Cardiol* **76**:55-70.
121. **Nakaya HI, Amaral PP, Louro R, Lopes A, Fachel AA, Moreira YB, El-Jundi TA, da Silva AM, Reis EM, Verjovski-Almeida S.** 2007. Genome mapping and expression analyses of human intronic noncoding RNAs reveal tissue-specific patterns and enrichment in genes related to regulation of transcription. *Genome Biol* **8**:R43.
 122. **Dinger ME, Pang KC, Mercer TR, Mattick JS.** 2008. Differentiating protein-coding and noncoding RNA: challenges and ambiguities. *PLoS Comput Biol* **4**:e1000176.
 123. **Chooniedass-Kothari S, Emberley E, Hamedani MK, Troup S, Wang X, Czosnek A, Hube F, Mutawe M, Watson PH, Leygue E.** 2004. The steroid receptor RNA activator is the first functional RNA encoding a protein. *FEBS Lett* **566**:43-47.
 124. **Warden CD, Kim SH, Yi SV.** 2008. Predicted functional RNAs within coding regions constrain evolutionary rates of yeast proteins. *PLoS One* **3**:e1559.
 125. **Jenny A, Hachet O, Zavorszky P, Cyrklaff A, Weston MD, Johnston DS, Erdelyi M, Ephrussi A.** 2006. A translation-independent role of oskar RNA in early *Drosophila* oogenesis. *Development* **133**:2827-2833.
 126. **Eastman G, Smircich P, Sotelo-Silveira JR.** 2018. Following Ribosome Footprints to Understand Translation at a Genome Wide Level. *Comput Struct Biotechnol J* **16**:167-176.
 127. **Guttman M, Russell P, Ingolia NT, Weissman JS, Lander ES.** 2013. Ribosome profiling provides evidence that large noncoding RNAs do not encode proteins. *Cell* **154**:240-251.
 128. **Wang L, Park HJ, Dasari S, Wang S, Kocher JP, Li W.** 2013. CPAT: Coding-Potential Assessment Tool using an alignment-free logistic regression model. *Nucleic Acids Res* **41**:e74.
 129. **Derrien T, Johnson R, Bussotti G, Tanzer A, Djebali S, Tilgner H, Guernec G, Martin D, Merkel A, Knowles DG, Lagarde J, Veeravalli L, Ruan X, Ruan Y, Lassmann T, Carninci P, Brown JB, Lipovich L, Gonzalez JM, Thomas M, Davis CA, Shiekhhattar R, Gingeras TR, Hubbard TJ, Notredame C, Harrow J, Guigo R.** 2012. The GENCODE v7 catalog of human long noncoding RNAs: analysis of their gene structure, evolution, and expression. *Genome Res* **22**:1775-1789.
 130. **Cabili MN, Trapnell C, Goff L, Koziol M, Tazon-Vega B, Regev A, Rinn JL.** 2011. Integrative annotation of human large intergenic noncoding RNAs reveals global properties and specific subclasses. *Genes Dev* **25**:1915-1927.
 131. **Djebali S, Davis CA, Merkel A, Dobin A, Lassmann T, Mortazavi A, Tanzer A, Lagarde J, Lin W, Schlesinger F, Xue C, Marinov GK, Khatun J, Williams BA, Zaleski C, Rozowsky J, Roder M, Kokocinski F, Abdelhamid RF, Alioto T, Antoshechkin I, Baer MT, Bar NS, Batut P, Bell K, Bell I, Chakraborty S, Chen X, Chrast J, Curado J, Derrien T, Drenkow J, Dumais E, Dumais J, Duttagupta R, Falconnet E, Fastuca M, Fejes-Toth K, Ferreira P, Foissac S, Fullwood MJ, Gao H, Gonzalez D, Gordon A, Gunawardena H, Howald C, Jha S, Johnson R, Kapranov P, King B, et al.** 2012. Landscape of transcription in human cells. *Nature* **489**:101-108.
 132. **Ravasi T, Suzuki H, Pang KC, Katayama S, Furuno M, Okunishi R, Fukuda S, Ru K, Frith MC, Gongora MM, Grimmond SM, Hume DA, Hayashizaki Y, Mattick JS.** 2006. Experimental validation of the regulated expression of large numbers of non-coding RNAs from the mouse genome. *Genome Res* **16**:11-19.
 133. **Cabanski CR, White NM, Dang HX, Silva-Fisher JM, Rauck CE, Cicka D, Maher CA.** 2015. Pan-cancer transcriptome analysis reveals long noncoding RNAs with conserved function. *RNA Biol* **12**:628-642.
 134. **Johnsson P, Lipovich L, Grandér D, Morris KV.** 2014. Evolutionary conservation of long non-coding RNAs; sequence, structure, function. *Biochim Biophys Acta* **1840**:1063-1071.
 135. **Lubelsky Y, Ulitsky I.** 2018. Sequences enriched in Alu repeats drive nuclear localization of long RNAs in human cells. *Nature* **555**:107-111.

136. **Li R, Zhu H, Luo Y.** 2016. Understanding the Functions of Long Non-Coding RNAs through Their Higher-Order Structures. *Int J Mol Sci* **17**.
137. **Xue Z, Hennelly S, Doyle B, Gulati AA, Novikova IV, Sanbonmatsu KY, Boyer LA.** 2016. A G-Rich Motif in the lncRNA Braveheart Interacts with a Zinc-Finger Transcription Factor to Specify the Cardiovascular Lineage. *Mol Cell* **64**:37-50.
138. **Li J, Ma W, Zeng P, Wang J, Geng B, Yang J, Cui Q.** 2015. LncTar: a tool for predicting the RNA targets of long noncoding RNAs. *Brief Bioinform* **16**:806-812.
139. **Lyle R, Watanabe D, te Vrugte D, Lerchner W, Smrzka OW, Wutz A, Schageman J, Hahner L, Davies C, Barlow DP.** 2000. The imprinted antisense RNA at the Igf2r locus overlaps but does not imprint Mas1. *Nat Genet* **25**:19-21.
140. **Su M, Wang H, Wang W, Wang Y, Ouyang L, Pan C, Xia L, Cao D, Liao Q.** 2018. LncRNAs in DNA damage response and repair in cancer cells. *Acta Biochim Biophys Sin (Shanghai)* **50**:433-439.
141. **Jiang Y, Li Y, Fang S, Jiang B, Qin C, Xie P, Zhou G, Li G.** 2014. The role of MALAT1 correlates with HPV in cervical cancer. *Oncol Lett* **7**:2135-2141.
142. **Bierhoff H, Dammert MA, Brocks D, Dambacher S, Schotta G, Grummt I.** 2014. Quiescence-induced LncRNAs trigger H4K20 trimethylation and transcriptional silencing. *Mol Cell* **54**:675-682.
143. **Clark MB, Johnston RL, Inostroza-Ponta M, Fox AH, Fortini E, Moscato P, Dinger ME, Mattick JS.** 2012. Genome-wide analysis of long noncoding RNA stability. *Genome Res* **22**:885-898.
144. **Atkinson SR, Marguerat S, Bitton DA, Rodriguez-Lopez M, Rallis C, Lemay JF, Cotobal C, Malecki M, Smialowski P, Mata J, Korber P, Bachand F, Bahler J.** 2018. Long noncoding RNA repertoire and targeting by nuclear exosome, cytoplasmic exonuclease, and RNAi in fission yeast. *Rna* **24**:1195-1213.
145. **Palazzo AF, Lee ES.** 2018. Sequence Determinants for Nuclear Retention and Cytoplasmic Export of mRNAs and lncRNAs. *Front Genet* **9**:440.
146. **Prasanth KV, Prasanth SG, Xuan Z, Hearn S, Freier SM, Bennett CF, Zhang MQ, Spector DL.** 2005. Regulating gene expression through RNA nuclear retention. *Cell* **123**:249-263.
147. **Hutchinson JN, Ensminger AW, Clemson CM, Lynch CR, Lawrence JB, Chess A.** 2007. A screen for nuclear transcripts identifies two linked noncoding RNAs associated with SC35 splicing domains. *BMC Genomics* **8**:39.
148. **Zhang K, Daigle JG, Cunningham KM, Coyne AN, Ruan K, Grima JC, Bowen KE, Wadhwa H, Yang P, Rigo F, Taylor JP, Gitler AD, Rothstein JD, Lloyd TE.** 2018. Stress Granule Assembly Disrupts Nucleocytoplasmic Transport. *Cell* **173**:958-971.e917.
149. **Amaral PP, Mattick JS.** 2008. Noncoding RNA in development. *Mamm Genome* **19**:454-492.
150. **Esteller M.** 2011. Non-coding RNAs in human disease. *Nat Rev Genet* **12**:861-874.
151. **Feng J, Bi C, Clark BS, Mady R, Shah P, Kohtz JD.** 2006. The Evf-2 noncoding RNA is transcribed from the Dlx-5/6 ultraconserved region and functions as a Dlx-2 transcriptional coactivator. *Genes Dev* **20**:1470-1484.
152. **Miao Y, Ajami NE, Huang TS, Lin FM, Lou CH, Wang YT, Li S, Kang J, Munkacsı H, Maurya MR, Gupta S, Chien S, Subramaniam S, Chen Z.** 2018. Enhancer-associated long non-coding RNA LEENE regulates endothelial nitric oxide synthase and endothelial function. *Nat Commun* **9**:292.
153. **Nagano T, Mitchell JA, Sanz LA, Pauler FM, Ferguson-Smith AC, Feil R, Fraser P.** 2008. The Air noncoding RNA epigenetically silences transcription by targeting G9a to chromatin. *Science* **322**:1717-1720.
154. **Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY.** 2010. Long noncoding RNA as modular scaffold of histone modification complexes. *Science* **329**:689-693.

155. **Hung T, Wang Y, Lin MF, Koegel AK, Kotake Y, Grant GD, Horlings HM, Shah N, Umbricht C, Wang P, Wang Y, Kong B, Langerod A, Borresen-Dale AL, Kim SK, van de Vijver M, Sukumar S, Whitfield ML, Kellis M, Xiong Y, Wong DJ, Chang HY.** 2011. Extensive and coordinated transcription of noncoding RNAs within cell-cycle promoters. *Nat Genet* **43**:621-629.
156. **Romero-Barrios N, Legascue MF, Benhamed M, Ariel F, Crespi M.** 2018. Splicing regulation by long noncoding RNAs. *Nucleic Acids Res* **46**:2169-2184.
157. **Tripathi V, Ellis JD, Shen Z, Song DY, Pan Q, Watt AT, Freier SM, Bennett CF, Sharma A, Bubulya PA, Blencowe BJ, Prasanth SG, Prasanth KV.** 2010. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell* **39**:925-938.
158. **Cao L, Zhang P, Li J, Wu M.** 2017. LAST, a c-Myc-inducible long noncoding RNA, cooperates with CNBP to promote CCND1 mRNA stability in human cells. *Elife* **6**.
159. **Yoon JH, Abdelmohsen K, Srikantan S, Yang X, Martindale JL, De S, Huarte M, Zhan M, Becker KG, Gorospe M.** 2012. LincRNA-p21 suppresses target mRNA translation. *Mol Cell* **47**:648-655.
160. **Militello G, Weirick T, John D, Doring C, Dimmeler S, Uchida S.** 2017. Screening and validation of lncRNAs and circRNAs as miRNA sponges. *Brief Bioinform* **18**:780-788.
161. **Olgun G, Sahin O, Tastan O.** 2018. Discovering lncRNA mediated sponge interactions in breast cancer molecular subtypes. *BMC Genomics* **19**:650.
162. **Agarwal V, Bell GW, Nam JW, Bartel DP.** 2015. Predicting effective microRNA target sites in mammalian mRNAs. *Elife* **4**.
163. **Furio-Tari P, Tarazona S, Gabaldon T, Enright AJ, Conesa A.** 2016. spongeScan: A web for detecting microRNA binding elements in lncRNA sequences. *Nucleic Acids Res* **44**:W176-180.
164. **Zhang A, Xu M, Mo YY.** 2014. Role of the lncRNA-p53 regulatory network in cancer. *J Mol Cell Biol* **6**:181-191.
165. **Li XL, Subramanian M, Jones MF, Chaudhary R, Singh DK, Zong X, Gryder B, Sindri S, Mo M, Schetter A, Wen X, Parvathaneni S, Kazandjian D, Jenkins LM, Tang W, Elloumi F, Martindale JL, Huarte M, Zhu Y, Robles AI, Frier SM, Rigo F, Cam M, Ambs S, Sharma S, Harris CC, Dasso M, Prasanth KV, Lal A.** 2017. Long Noncoding RNA PURPL Suppresses Basal p53 Levels and Promotes Tumorigenicity in Colorectal Cancer. *Cell Rep* **20**:2408-2423.
166. **Guo JU, Agarwal V, Guo H, Bartel DP.** 2014. Expanded identification and characterization of mammalian circular RNAs. *Genome Biol* **15**:409.
167. **Conn SJ, Pillman KA, Toubia J, Conn VM, Salmanidis M, Phillips CA, Roslan S, Schreiber AW, Gregory PA, Goodall GJ.** 2015. The RNA binding protein quaking regulates formation of circRNAs. *Cell* **160**:1125-1134.
168. **Li S, Teng S, Xu J, Su G, Zhang Y, Zhao J, Zhang S, Wang H, Qin W, Lu ZJ, Guo Y, Zhu Q, Wang D.** 2018. Microarray is an efficient tool for circRNA profiling. *Brief Bioinform* doi:10.1093/bib/bby006.
169. **Hansen TB, Kjems J, Damgaard CK.** 2013. Circular RNA and miR-7 in cancer. *Cancer Res* **73**:5609-5612.
170. **Hansen TB, Jensen TI, Clausen BH, Bramsen JB, Finsen B, Damgaard CK, Kjems J.** 2013. Natural RNA circles function as efficient microRNA sponges. *Nature* **495**:384-388.
171. **Du WW, Zhang C, Yang W, Yong T, Awan FM, Yang BB.** 2017. Identifying and Characterizing circRNA-Protein Interaction. *Theranostics* **7**:4183-4191.
172. **Ashwal-Fluss R, Meyer M, Pamudurti NR, Ivanov A, Bartok O, Hanan M, Evantal N, Memczak S, Rajewsky N, Kadener S.** 2014. circRNA biogenesis competes with pre-mRNA splicing. *Mol Cell* **56**:55-66.
173. **Pamudurti NR, Bartok O, Jens M, Ashwal-Fluss R, Stottmeister C, Ruhe L, Hanan M, Wyler E, Perez-Hernandez D, Ramberger E, Shenzen S, Samson M, Dittmar G, Landthaler M, Chekulaeva M, Rajewsky N, Kadener S.** 2017. Translation of CircRNAs. *Mol Cell* **66**:9-21.e27.

174. **Yang Y, Fan X, Mao M, Song X, Wu P, Zhang Y, Jin Y, Yang Y, Chen LL, Wang Y, Wong CC, Xiao X, Wang Z.** 2017. Extensive translation of circular RNAs driven by N(6)-methyladenosine. *Cell Res* **27**:626-641.
175. **Ungerleider N, Concha M, Lin Z, Roberts C, Wang X, Cao S, Baddoo M, Moss WN, Yu Y, Seddon M, Lehman T, Tibbetts S, Renne R, Dong Y, Flemington EK.** 2018. The Epstein Barr virus circRNAome. *PLoS Pathog* **14**:e1007206.
176. **Tagawa T, Gao S, Koparde VN, Gonzalez M, Spouge JL, Serquina AP, Lurain K, Ramaswami R, Uldrick TS, Yarchoan R, Ziegelbauer JM.** 2018. Discovery of Kaposi's sarcoma herpesvirus-encoded circular RNAs and a human antiviral circular RNA. *Proc Natl Acad Sci U S A* doi:10.1073/pnas.1816183115.
177. **Zhong ZJ, Yang JX, Cao DY, Sun Y, Sun LL, Cheng XM, Chen J, Lang JH, Shen K.** 2012. [Promoter methylation of DAPK1, RAR-beta and MGMT in exfoliated cervical cytology and its clinical application]. *Zhonghua Fu Chan Ke Za Zhi* **47**:196-200.
178. **Shivapurkar N, Sherman ME, Stastny V, Echebiri C, Rader JS, Nayar R, Bonfiglio TA, Gazdar AF, Wang SS.** 2007. Evaluation of candidate methylation markers to detect cervical neoplasia. *Gynecol Oncol* **107**:549-553.
179. **Lim EH, Ng SL, Li JL, Chang AR, Ng J, Ilancheran A, Low J, Quek SC, Tay EH.** 2010. Cervical dysplasia: assessing methylation status (Methylight) of CCNA1, DAPK1, HS3ST2, PAX1 and TFPI2 to improve diagnostic accuracy. *Gynecol Oncol* **119**:225-231.
180. **Bierkens M, Hesselink AT, Meijer CJ, Heideman DA, Wisman GB, van der Zee AG, Snijders PJ, Steenbergen RD.** 2013. CADM1 and MAL promoter methylation levels in hrHPV-positive cervical scrapes increase proportional to degree and duration of underlying cervical disease. *Int J Cancer* **133**:1293-1299.
181. **Fang J, Zhang H, Jin S.** 2014. Epigenetics and cervical cancer: from pathogenesis to therapy. *Tumour Biol* **35**:5083-5093.
182. **Parasramka MA, Maji S, Matsuda A, Yan IK, Patel T.** 2016. Long non-coding RNAs as novel targets for therapy in hepatocellular carcinoma. *Pharmacol Ther* **161**:67-78.
183. **Arun G, Diermeier SD, Spector DL.** 2018. Therapeutic Targeting of Long Non-Coding RNAs in Cancer. *Trends Mol Med* **24**:257-277.
184. **Gupta RA, Shah N, Wang KC, Kim J, Horlings HM, Wong DJ, Tsai MC, Hung T, Argani P, Rinn JL, Wang Y, Brzoska P, Kong B, Li R, West RB, van de Vijver MJ, Sukumar S, Chang HY.** 2010. Long non-coding RNA HOTAIR reprograms chromatin state to promote cancer metastasis. *Nature* **464**:1071-1076.
185. **Mao X, Qin X, Li L, Zhou J, Zhou M, Li X, Xu Y, Yuan L, Liu QN, Xing H.** 2018. A 15-long non-coding RNA signature to improve prognosis prediction of cervical squamous cell carcinoma. *Gynecol Oncol* doi:10.1016/j.ygyno.2017.12.011.
186. **Yang JP, Yang XJ, Xiao L, Wang Y.** 2016. Long noncoding RNA PVT1 as a novel serum biomarker for detection of cervical cancer. *Eur Rev Med Pharmacol Sci* **20**:3980-3986.
187. **Liang W, Lv T, Shi X, Liu H, Zhu Q, Zeng J, Yang W, Yin J, Song Y.** 2016. Circulating long noncoding RNA GAS5 is a novel biomarker for the diagnosis of nonsmall cell lung cancer. *Medicine (Baltimore)* **95**:e4608.
188. **Durst M, Gissmann L, Ikenberg H, zur Hausen H.** 1983. A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. *Proc Natl Acad Sci U S A* **80**:3812-3815.
189. **Moody C.** 2017. Mechanisms by which HPV Induces a Replication Competent Environment in Differentiating Keratinocytes. *Viruses* **9**.
190. **Halbert CL, Demers GW, Galloway DA.** 1991. The E7 gene of human papillomavirus type 16 is sufficient for immortalization of human epithelial cells. *J Virol* **65**:473-478.
191. **Moody CA, Laimins LA.** 2010. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer* **10**:550-560.

192. **de Martel C, Plummer M, Vignat J, Franceschi S.** 2017. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int J Cancer* **141**:664-670.
193. **Harper DM, DeMars LR.** 2017. HPV vaccines - A review of the first decade. *Gynecol Oncol* **146**:196-204.
194. **Tao M, Kruhlak M, Xia S, Androphy E, Zheng ZM.** 2003. Signals that dictate nuclear localization of human papillomavirus type 16 oncoprotein E6 in living cells. *J Virol* **77**:13232-13247.
195. **Jabbar S, Strati K, Shin MK, Pitot HC, Lambert PF.** 2010. Human papillomavirus type 16 E6 and E7 oncoproteins act synergistically to cause head and neck cancer in mice. *Virology* **407**:60-67.
196. **Mesplede T, Gagnon D, Bergeron-Labrecque F, Azar I, Senechal H, Coutlee F, Archambault J.** 2012. p53 degradation activity, expression, and subcellular localization of E6 proteins from 29 human papillomavirus genotypes. *J Virol* **86**:94-107.
197. **Yeung CL, Tsang TY, Yau PL, Kwok TT.** 2017. Human papillomavirus type 16 E6 suppresses microRNA-23b expression in human cervical cancer cells through DNA methylation of the host gene C9orf3. *Oncotarget* **8**:12158-12173.
198. **Burgers WA, Blanchon L, Pradhan S, de Launoit Y, Kouzarides T, Fuks F.** 2007. Viral oncoproteins target the DNA methyltransferases. *Oncogene* **26**:1650-1655.
199. **Laurson J, Khan S, Chung R, Cross K, Raj K.** 2010. Epigenetic repression of E-cadherin by human papillomavirus 16 E7 protein. *Carcinogenesis* **31**:918-926.
200. **D'Costa ZJ, Jolly C, Androphy EJ, Mercer A, Matthews CM, Hibma MH.** 2012. Transcriptional repression of E-cadherin by human papillomavirus type 16 E6. *PLoS One* **7**:e48954.
201. **McCabe MT, Davis JN, Day ML.** 2005. Regulation of DNA methyltransferase 1 by the pRb/E2F1 pathway. *Cancer Res* **65**:3624-3632.
202. **Scheffner M, Huibregtse JM, Vierstra RD, Howley PM.** 1993. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* **75**:495-505.
203. **Zimmermann H, Degenkolbe R, Bernard HU, O'Connor MJ.** 1999. The human papillomavirus type 16 E6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator CBP/p300. *J Virol* **73**:6209-6219.
204. **Patel D, Huang SM, Baglia LA, McCance DJ.** 1999. The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. *Embo j* **18**:5061-5072.
205. **Ojesina AI, Lichtenstein L, Freeman SS, Peadarallu CS, Imaz-Rosshandler I, Pugh TJ, Cherniack AD, Ambrogio L, Cibulskis K, Bertelsen B, Romero-Cordoba S, Trevino V, Vazquez-Santillan K, Guadarrama AS, Wright AA, Rosenberg MW, Duke F, Kaplan B, Wang R, Nickerson E, Walline HM, Lawrence MS, Stewart C, Carter SL, McKenna A, Rodriguez-Sanchez IP, Espinosa-Castilla M, Woie K, Bjorge L, Wik E, Halle MK, Hoivik EA, Krakstad C, Gabino NB, Gomez-Macias GS, Valdez-Chapa LD, Garza-Rodriguez ML, Maytorena G, Vazquez J, Rodea C, Cravioto A, Cortes ML, Greulich H, Crum CP, Neuberger DS, Hidalgo-Miranda A, Escareno CR, Akslen LA, Carey TE, Vintermyr OK, et al.** 2014. Landscape of genomic alterations in cervical carcinomas. *Nature* **506**:371-375.
206. **Gibb EA, Becker-Santos DD, Enfield KS, Guillaud M, Niekerk D, Matisic JP, Macaulay CE, Lam WL.** 2012. Aberrant expression of long noncoding RNAs in cervical intraepithelial neoplasia. *Int J Gynecol Cancer* **22**:1557-1563.
207. **Wang H, Zhao Y, Chen M, Cui J.** 2017. Identification of Novel Long Non-coding and Circular RNAs in Human Papillomavirus-Mediated Cervical Cancer. *Front Microbiol* **8**:1720.
208. **Yu D, Li Y, Ming Z, Wang H, Dong Z, Qiu L, Wang T.** 2018. Comprehensive circular RNA expression profile in radiation-treated HeLa cells and analysis of radioresistance-related circRNAs. *PeerJ* **6**:e5011.
209. **Dong J, Su M, Chang W, Zhang K, Wu S, Xu T.** 2017. Long non-coding RNAs on the stage of cervical cancer (Review). *Oncol Rep* **38**:1923-1931.

210. **Aalijahan H, Ghorbian S.** 2018. Long non-coding RNAs and cervical cancer. *Exp Mol Pathol* doi:10.1016/j.yexmp.2018.11.010.
211. **Kim HJ, Lee DW, Yim GW, Nam EJ, Kim S, Kim SW, Kim YT.** 2015. Long non-coding RNA HOTAIR is associated with human cervical cancer progression. *Int J Oncol* **46**:521-530.
212. **Huang L, Liao LM, Liu AW, Wu JB, Cheng XL, Lin JX, Zheng M.** 2014. Overexpression of long noncoding RNA HOTAIR predicts a poor prognosis in patients with cervical cancer. *Arch Gynecol Obstet* **290**:717-723.
213. **Hajjari M, Salavaty A.** 2015. HOTAIR: an oncogenic long non-coding RNA in different cancers. *Cancer Biol Med* **12**:1-9.
214. **Jing L, Yuan W, Ruofan D, Jinjin Y, Haifeng Q.** 2015. HOTAIR enhanced aggressive biological behaviors and induced radio-resistance via inhibiting p21 in cervical cancer. *Tumour Biol* **36**:3611-3619.
215. **Gutschner T, Hammerle M, Diederichs S.** 2013. MALAT1 -- a paradigm for long noncoding RNA function in cancer. *J Mol Med (Berl)* **91**:791-801.
216. **Guo F, Li Y, Liu Y, Wang J, Li Y, Li G.** 2010. Inhibition of metastasis-associated lung adenocarcinoma transcript 1 in CaSki human cervical cancer cells suppresses cell proliferation and invasion. *Acta Biochim Biophys Sin (Shanghai)* **42**:224-229.
217. **Yang L, Bai HS, Deng Y, Fan L.** 2015. High MALAT1 expression predicts a poor prognosis of cervical cancer and promotes cancer cell growth and invasion. *Eur Rev Med Pharmacol Sci* **19**:3187-3193.
218. **Liu S, Song L, Zeng S, Zhang L.** 2016. MALAT1-miR-124-RBG2 axis is involved in growth and invasion of HR-HPV-positive cervical cancer cells. *Tumour Biol* **37**:633-640.
219. **Sun NX, Ye C, Zhao Q, Zhang Q, Xu C, Wang SB, Jin ZJ, Sun SH, Wang F, Li W.** 2014. Long noncoding RNA-EBIC promotes tumor cell invasion by binding to EZH2 and repressing E-cadherin in cervical cancer. *PLoS One* **9**:e100340.
220. **Zhang D, Sun G, Zhang H, Tian J, Li Y.** 2017. Long non-coding RNA ANRIL indicates a poor prognosis of cervical cancer and promotes carcinogenesis via PI3K/Akt pathways. *Biomed Pharmacother* **85**:511-516.
221. **Zhang JJ, Wang DD, Du CX, Wang Y.** 2018. Long Noncoding RNA ANRIL Promotes Cervical Cancer Development by Acting as a Sponge of miR-186. *Oncol Res* **26**:345-352.
222. **Zhu H, Chen X, Hu Y, Shi Z, Zhou Q, Zheng J, Wang Y.** 2017. Long non-coding RNA expression profile in cervical cancer tissues. *Oncol Lett* **14**:1379-1386.
223. **Chen J, Fu Z, Ji C, Gu P, Xu P, Yu N, Kan Y, Wu X, Shen R, Shen Y.** 2015. Systematic gene microarray analysis of the lncRNA expression profiles in human uterine cervix carcinoma. *Biomed Pharmacother* **72**:83-90.
224. **Liu Y, Yang Y, Li L, Liu Y, Geng P, Li G, Song H.** 2018. LncRNA SNHG1 enhances cell proliferation, migration, and invasion in cervical cancer. *Biochem Cell Biol* **96**:38-43.
225. **Huang L, Jiang X, Wang Z, Zhong X, Tai S, Cui Y.** 2018. Small nucleolar RNA host gene 1: A new biomarker and therapeutic target for cancers. *Pathol Res Pract* **214**:1247-1252.
226. **Chen Y, Wang CX, Sun XX, Wang C, Liu TF, Wang DJ.** 2017. Long non-coding RNA CCHE1 overexpression predicts a poor prognosis for cervical cancer. *Eur Rev Med Pharmacol Sci* **21**:479-483.
227. **Yang M, Zhai X, Xia B, Wang Y, Lou G.** 2015. Long noncoding RNA CCHE1 promotes cervical cancer cell proliferation via upregulating PCNA. *Tumour Biol* **36**:7615-7622.
228. **Graham LD, Pedersen SK, Brown GS, Ho T, Kassir Z, Moynihan AT, Vizgoft EK, Dunne R, Pimlott L, Young GP, Lapointe LC, Molloy PL.** 2011. Colorectal Neoplasia Differentially Expressed (CRNDE), a Novel Gene with Elevated Expression in Colorectal Adenomas and Adenocarcinomas. *Genes Cancer* **2**:829-840.
229. **Wang Y, Wang Y, Li J, Zhang Y, Yin H, Han B.** 2015. CRNDE, a long-noncoding RNA, promotes glioma cell growth and invasion through mTOR signaling. *Cancer Lett* **367**:122-128.

230. **Xie H, Ma B, Gao Q, Zhan H, Liu Y, Chen Z, Ye S, Li J, Yao L, Huang W.** 2018. Long non-coding RNA CRNDE in cancer prognosis: Review and meta-analysis. *Clin Chim Acta* **485**:262-271.
231. **Meng Y, Li Q, Li L, Ma R.** 2017. The long non-coding RNA CRNDE promotes cervical cancer cell growth and metastasis. *Biol Chem* **399**:93-100.
232. **Yang HY, Huang CP, Cao MM, Wang YF, Liu Y.** 2018. Long non-coding RNA CRNDE may be associated with poor prognosis by promoting proliferation and inhibiting apoptosis of cervical cancer cells through targeting PI3K/AKT. *Neoplasma* doi:10.4149/neo_2018_171225N841.
233. **Qin R, Chen Z, Ding Y, Hao J, Hu J, Guo F.** 2013. Long non-coding RNA MEG3 inhibits the proliferation of cervical carcinoma cells through the induction of cell cycle arrest and apoptosis. *Neoplasma* **60**:486-492.
234. **Zhang J, Yao T, Wang Y, Yu J, Liu Y, Lin Z.** 2016. Long noncoding RNA MEG3 is downregulated in cervical cancer and affects cell proliferation and apoptosis by regulating miR-21. *Cancer Biol Ther* **17**:104-113.
235. **Godinho MF, Wulfschlegel JD, Look MP, Sieuwerts AM, Sleijfer S, Foekens JA, Petricoin EF, 3rd, Dorssers LC, van Agthoven T.** 2012. BCAR4 induces antioestrogen resistance but sensitises breast cancer to lapatinib. *Br J Cancer* **107**:947-955.
236. **Anonymous.** 2017. Integrated genomic and molecular characterization of cervical cancer. *Nature* **543**:378-384.
237. **Corcoran DL, Pandit KV, Gordon B, Bhattacharjee A, Kaminski N, Benos PV.** 2009. Features of mammalian microRNA promoters emerge from polymerase II chromatin immunoprecipitation data. *PLoS One* **4**:e5279.
238. **Shi Y, Lu J, Zhou J, Tan X, He Y, Ding J, Tian Y, Wang L, Wang K.** 2014. Long non-coding RNA Loc554202 regulates proliferation and migration in breast cancer cells. *Biochem Biophys Res Commun* **446**:448-453.
239. **Yang L, Wei H, Xiao HJ.** 2016. Long non-coding RNA Loc554202 expression as a prognostic factor in patients with colorectal cancer. *Eur Rev Med Pharmacol Sci* **20**:4243-4247.
240. **Chen J, Zhu J.** 2017. Elevated Expression Levels of Long Non-Coding RNA, Loc554202, Are Predictive of Poor Prognosis in Cervical Cancer. *Tohoku J Exp Med* **243**:165-172.
241. **Shima H, Kida K, Adachi S, Yamada A, Sugae S, Narui K, Miyagi Y, Nishi M, Ryo A, Murata S, Taniguchi H, Ichikawa Y, Ishikawa T, Endo I.** 2018. Lnc RNA H19 is associated with poor prognosis in breast cancer patients and promotes cancer stemness. *Breast Cancer Res Treat* doi:10.1007/s10549-018-4793-z.
242. **Zhu M, Chen Q, Liu X, Sun Q, Zhao X, Deng R, Wang Y, Huang J, Xu M, Yan J, Yu J.** 2014. lncRNA H19/miR-675 axis represses prostate cancer metastasis by targeting TGFBI. *FEBS J* **281**:3766-3775.
243. **Iempridee T.** 2017. Long non-coding RNA H19 enhances cell proliferation and anchorage-independent growth of cervical cancer cell lines. *Exp Biol Med (Maywood)* **242**:184-193.
244. **Ou L, Wang D, Zhang H, Yu Q, Hua F.** 2018. Decreased Expression of miR-138-5p by lncRNA H19 in Cervical Cancer Promotes Tumor Proliferation. *Oncol Res* **26**:401-410.
245. **Mourtada-Maarabouni M, Pickard MR, Hedge VL, Farzaneh F, Williams GT.** 2009. GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene* **28**:195-208.
246. **Shi X, Sun M, Liu H, Yao Y, Kong R, Chen F, Song Y.** 2015. A critical role for the long non-coding RNA GAS5 in proliferation and apoptosis in non-small-cell lung cancer. *Mol Carcinog* **54 Suppl 1**:E1-e12.
247. **Ma C, Shi X, Zhu Q, Li Q, Liu Y, Yao Y, Song Y.** 2016. The growth arrest-specific transcript 5 (GAS5): a pivotal tumor suppressor long noncoding RNA in human cancers. *Tumour Biol* **37**:1437-1444.
248. **Cao S, Liu W, Li F, Zhao W, Qin C.** 2014. Decreased expression of lncRNA GAS5 predicts a poor prognosis in cervical cancer. *Int J Clin Exp Pathol* **7**:6776-6783.

249. **Liao LM, Sun XY, Liu AW, Wu JB, Cheng XL, Lin JX, Zheng M, Huang L.** 2014. Low expression of long noncoding XLOC_010588 indicates a poor prognosis and promotes proliferation through upregulation of c-Myc in cervical cancer. *Gynecol Oncol* **133**:616-623.
250. **Jiang S, Wang HL, Yang J.** 2015. Low expression of long non-coding RNA LET inhibits carcinogenesis of cervical cancer. *Int J Clin Exp Pathol* **8**:806-811.
251. **Iempridee T, Wiwithaphon S, Piboonprai K, Pratedrat P, Khumkhong P, Japrun D, Temisak S, Laiwejpithaya S, Chaopotong P, Dharakul T.** 2018. Identification of reference genes for circulating long noncoding RNA analysis in serum of cervical cancer patients. *FEBS Open Bio* **8**:1844-1854.
252. **Gao YL, Zhang MY, Xu B, Han LJ, Lan SF, Chen J, Dong YJ, Cao LL.** 2017. Circular RNA expression profiles reveal that hsa_circ_0018289 is up-regulated in cervical cancer and promotes the tumorigenesis. *Oncotarget* **8**:86625-86633.
253. **Zhang J, Zhao X, Zhang J, Zheng X, Li F.** 2018. Circular RNA hsa_circ_0023404 exerts an oncogenic role in cervical cancer through regulating miR-136/TFCP2/YAP pathway. *Biochem Biophys Res Commun* **501**:428-433.
254. **Ma HB, Yao YN, Yu JJ, Chen XX, Li HF.** 2018. Extensive profiling of circular RNAs and the potential regulatory role of circRNA-000284 in cell proliferation and invasion of cervical cancer via sponging miR-506. *Am J Transl Res* **10**:592-604.
255. **Liu J, Wang D, Long Z, Liu J, Li W.** 2018. CircRNA8924 Promotes Cervical Cancer Cell Proliferation, Migration and Invasion by Competitively Binding to MiR-518d-5p /519-5p Family and Modulating the Expression of CBX8. *Cell Physiol Biochem* **48**:173-184.
256. **Li Z, Huang C, Bao C, Chen L, Lin M, Wang X, Zhong G, Yu B, Hu W, Dai L, Zhu P, Chang Z, Wu Q, Zhao Y, Jia Y, Xu P, Liu H, Shan G.** 2015. Exon-intron circular RNAs regulate transcription in the nucleus. *Nat Struct Mol Biol* **22**:256-264.
257. **Tomar S, Graves CA, Altomare D, Kowli S, Kassler S, Sutkowski N, Gillespie MB, Creek KE, Pirisi L.** 2016. Human papillomavirus status and gene expression profiles of oropharyngeal and oral cancers from European American and African American patients. *Head Neck* **38 Suppl 1**:E694-704.
258. **Yang L, Yi K, Wang H, Zhao Y, Xi M.** 2016. Comprehensive analysis of lncRNAs microarray profile and mRNA-lncRNA co-expression in oncogenic HPV-positive cervical cancer cell lines. *Oncotarget* **7**:49917-49929.
259. **Nohata N, Abba MC, Gutkind JS.** 2016. Unraveling the oral cancer lncRNAome: Identification of novel lncRNAs associated with malignant progression and HPV infection. *Oral Oncol* **59**:58-66.
260. **Gyongyosi E, Szalmas A, Ferenczi A, Poliska S, Konya J, Veress G.** 2015. Transcriptional regulation of genes involved in keratinocyte differentiation by human papillomavirus 16 oncoproteins. *Arch Virol* **160**:389-398.
261. **Villota C, Campos A, Vidaurre S, Oliveira-Cruz L, Boccardo E, Burzio VA, Varas M, Villegas J, Villa LL, Valenzuela PD, Socias M, Roberts S, Burzio LO.** 2012. Expression of mitochondrial non-coding RNAs (ncRNAs) is modulated by high risk human papillomavirus (HPV) oncogenes. *J Biol Chem* **287**:21303-21315.
262. **Sharma S, Munger K.** 2018. Expression of the cervical carcinoma expressed PCNA regulatory (CCEPR) long noncoding RNA is driven by the human papillomavirus E6 protein and modulates cell proliferation independent of PCNA. *Virology* **518**:8-13.
263. **Iancu IV, Anton G, Botezatu A, Huica I, Nastase A, Socolov DG, Stanescu AD, Dima SO, Bacalbasa N, Plesa A.** 2017. LINC01101 and LINC00277 expression levels as novel factors in HPV-induced cervical neoplasia. *J Cell Mol Med* **21**:3787-3794.
264. **Zheng SR, Zhang HR, Zhang ZF, Lai SY, Huang LJ, Liu J, Bai X, Ding K, Zhou JY.** 2018. Human papillomavirus 16 E7 oncoprotein alters the expression profiles of circular RNAs in Caski cells. *J Cancer* **9**:3755-3764.

265. **Sharma S, Mandal P, Sadhukhan T, Roy Chowdhury R, Ranjan Mondal N, Chakravarty B, Chatterjee T, Roy S, Sengupta S.** 2015. Bridging Links between Long Noncoding RNA HOTAIR and HPV Oncoprotein E7 in Cervical Cancer Pathogenesis. *Sci Rep* **5**:11724.
266. **Liu S, Song L, Yao H, Zhang L, Xu D, Gao F, Li Q.** 2016. MiR-375 Is Epigenetically Downregulated by HPV-16 E6 Mediated DNMT1 Upregulation and Modulates EMT of Cervical Cancer Cells by Suppressing lncRNA MALAT1. *PLoS One* **11**:e0163460.
267. **Lu H, He Y, Lin L, Qi Z, Ma L, Li L, Su Y.** 2016. Long non-coding RNA MALAT1 modulates radiosensitivity of HR-HPV+ cervical cancer via sponging miR-145. *Tumour Biol* **37**:1683-1691.
268. **Jeffers LK, Duan K, Ellies LG, Seaman WT, Burger-Calderon RA, Diatchenko LB, Webster-Cyriaque J.** 2013. Correlation of transcription of MALAT-1, a novel noncoding RNA, with deregulated expression of tumor suppressor p53 in small DNA tumor virus models. *J Cancer Ther* **4**.
269. **White NM, Cabanski CR, Silva-Fisher JM, Dang HX, Govindan R, Maher CA.** 2014. Transcriptome sequencing reveals altered long intergenic non-coding RNAs in lung cancer. *Genome Biol* **15**:429.
270. **Yang F, Lv SX, Lv L, Liu YH, Dong SY, Yao ZH, Dai XX, Zhang XH, Wang OC.** 2016. Identification of lncRNA FAM83H-AS1 as a novel prognostic marker in luminal subtype breast cancer. *Onco Targets Ther* **9**:7039-7045.
271. **Yang L, Xu L, Wang Q, Wang M, An G.** 2016. Dysregulation of long non-coding RNA profiles in human colorectal cancer and its association with overall survival. *Oncol Lett* **12**:4068-4074.
272. **Zhang J, Feng S, Su W, Bai S, Xiao L, Wang L, Thomas DG, Lin J, Reddy RM, Carrott PW, Lynch WR, Chang AC, Beer DG, Guo YM, Chen G.** 2017. Overexpression of FAM83H-AS1 indicates poor patient survival and knockdown impairs cell proliferation and invasion via MET/EGFR signaling in lung cancer. *Sci Rep* **7**:42819.
273. **Baratieh Z, Khalaj Z, Honardoost MA, Emadi-Baygi M, Khanahmad H, Salehi M, Nikpour P.** 2017. Aberrant expression of PlncRNA-1 and TUG1: potential biomarkers for gastric cancer diagnosis and clinically monitoring cancer progression. *Biomark Med* **11**:1077-1090.
274. **Arnes L, Liu Z, Wang J, Carlo Maurer H, Sagalovskiy I, Sanchez-Martin M, Bommakanti N, Garofalo DC, Balderes DA, Sussel L, Olive KP, Rabadan R.** 2018. Comprehensive characterisation of compartment-specific long non-coding RNAs associated with pancreatic ductal adenocarcinoma. *Gut* doi:10.1136/gutjnl-2017-314353.
275. **Bi YY, Shen G, Quan Y, Jiang W, Xu F.** 2018. Long noncoding RNA FAM83H-AS1 exerts an oncogenic role in glioma through epigenetically silencing CDKN1A (p21). *J Cell Physiol* doi:10.1002/jcp.26813.
276. **Lu S, Dong W, Zhao P, Liu Z.** 2018. lncRNA FAM83H-AS1 is associated with the prognosis of colorectal carcinoma and promotes cell proliferation by targeting the Notch signaling pathway. *Oncol Lett* **15**:1861-1868.

Chapter 2

Long non-coding RNA FAM83H-AS1 is regulated by human papillomavirus 16 E6 independently of p53 in cervical cancer cells

Jamie A. Barr¹, Karen E. Hayes¹, Tayvia Brownmiller¹, Abby D. Harold¹, Rajaganapathi Jagannathan²,
Paul R. Lockman², Saleem Khan³, and Ivan Martinez^{1,*}

¹Department of Microbiology, West Virginia University Cancer Institute, School of Medicine, West Virginia University, Morgantown, West Virginia 26506; ²Department of Basic Pharmaceutical Sciences, Health Sciences Center, School of Pharmacy, West Virginia University, Morgantown, West Virginia 26506; ³Department of Microbiology and Molecular Genetics, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15219

Key Words: Long non-coding RNA, Human papillomavirus, FAM83H-AS1, Cervical cancer

***Correspondence:** Ivan Martinez, (ivmartinez@hsc.wvu.edu)

*The authors disclose no potential conflicts of interest.

Published in: Scientific Reports

Sci Rep. 2019 Mar 6;9(1):3662. PMID: 30842470

I. Abstract

High-risk human papillomavirus (HPV) infection is one of the first events in the process of carcinogenesis in cervical and head and neck cancers. The expression of the viral oncoproteins E6 and E7 are essential in this process by inactivating the tumor suppressor proteins p53 and Rb, respectively, in addition to their interactions with other host proteins. Non-coding RNAs, such as long non-coding RNAs (lncRNAs) have been found to be dysregulated in several cancers, suggesting an important role in tumorigenesis. In order to identify host lncRNAs affected by HPV infection, we expressed the high-risk HPV-16 E6 oncoprotein in primary human keratinocytes and measured the global lncRNA expression profile by high-throughput sequencing (RNA-seq). We found several host lncRNAs differentially expressed by E6 including GAS5, H19, and FAM83H-AS1. Interestingly, FAM83H-AS1 was found overexpressed in HPV-16 positive cervical cancer cell lines in an HPV-16 E6-dependent manner but independently of p53 regulation. Furthermore, FAM83H-AS1 was found to be regulated through E6-p300 pathway. Knockdown of FAM83H-AS1 by siRNAs decreased cellular proliferation, migration and increased apoptosis. FAM83H-AS1 was also found to be altered in human cervical cancer tissues and high expression of this lncRNA was associated with worse overall survival, suggesting an important role in cervical carcinogenesis.

II. Introduction

High-risk HPV infection (e.g. HPV-16) is one of the most common causes of cervical cancer(1-3), as well as a subset of head and neck squamous cell carcinoma (HNSCC)(1). The HPV oncoproteins E6 and E7 have been shown to contribute to carcinogenesis by modulating the degradation of human proteins, such as the tumor suppressors p53(4) and Rb(5) as well as a plethora of other cellular proteins(2, 3, 6-8). The HPV-16 E6 protein can abrogate p53 function by proteasomal degradation as it forms a complex with E6-associated protein (E6AP)(9), or by targeting the p53 coactivator CBP-p300(8, 10). Upon transmission, HPV infects the undifferentiated keratinocytes at the basal layer of the stratified epithelia and its genome remains episomal maintaining low copy numbers. During the course of cancer development, the viral genome frequently becomes integrated into the host cell DNA(11).

The recent discovery of different classes of non-coding RNAs (ncRNAs) expressed in human cells has opened a new chapter in the understanding of cellular processes, such as chromatin remodeling, transcriptional control, and post-transcriptional regulation. One of these classes of ncRNAs called long

non-coding RNAs (lncRNAs) are defined as RNAs larger than 200 nucleotides that are not translated into proteins. Recent findings indicate that lncRNAs are involved in gene regulation at the transcriptional level by functioning as signal, guide, decoy, or scaffold RNAs(12-14). Dysregulation of lncRNAs occurs in a variety of cancers, suggesting a potential use of these ncRNAs as biomarkers for diagnosis, prognosis, stage of cancer, and response to therapy(15-18). LncRNAs have been shown to be altered in cervical cancer(19-21), however, only a few publications have studied lncRNAs that are specifically regulated by the HPV E6 oncoprotein, such as MALAT1 and CCEPR(22, 23). These lncRNAs were found altered in cervical cancer but it is unknown if these alterations are part of the early events in cervical carcinogenesis. A few studies have looked at aberrant expression of lncRNAs in progression from pre-malignant cervical intraepithelial neoplasia (CIN) to cervical cancer(24, 25). For these reasons, it is important to understand if certain lncRNAs are important in the first stages of immortalization and transformation caused by HPV infections.

In this study, we demonstrated that the lncRNA FAM83H-AS1 (also known as onco-lncRNA-3) is up-regulated in primary keratinocytes expressing HPV-16 oncogene E6 as well as HPV-16 positive human cervical cancer cell lines and cervical tumor samples. We show that FAM83H-AS1 is regulated by HPV-16 E6 through the presence of p300 instead of the tumor suppressor p53. Finally, we show that FAM83H-AS1 is involved with cellular proliferation, migration, and apoptosis and is associated with worse overall survival in cervical cancer patients.

III. Results

A. High-risk HPV-16 E6 oncoprotein alters host long non-coding RNAs in primary keratinocytes.

As an initial screen to identify host lncRNAs that are regulated specifically by HPV-16 E6, we developed a system to look specifically at the effect of E6 expression alone in primary human foreskin keratinocytes (HEKa). HEKa were infected with a retroviral vector expressing HPV-16 E6 oncogene or GFP as a control. After puromycin selection and stable expression of HPV-16 E6 was confirmed by RT-PCR (Fig. S1), RNA was extracted, and samples were analyzed by RNA high-throughput sequencing (RNA-seq) to determine gene expression alterations in long non-coding RNAs (lncRNAs). Following bioinformatics analysis, we found 151 up- and 100 down-regulated host lncRNAs altered greater than 1.5-fold change when HPV-16 E6 was expressed in HEKa cells compared to GFP control (Fig. 1A, Table S1). From these host lncRNAs, we randomly chose 8 up- and 8 down-regulated host lncRNAs to validate our RNA-seq data by qRT-PCR. The expression of all the lncRNAs selected for validation followed the

same trend of up- or down-regulation found by RNA-seq (Fig. 1B). In order to determine the importance of these 251 lncRNAs in cervical carcinogenesis, a variety of filtering methods were utilized to reduce our scope (see Methods section). After filtering, we performed preliminary experiments with many lncRNAs, however, some of our top altered lncRNAs from our RNAseq analysis of foreskin keratinocytes (HEKa) were not altered in cervical cells (HCK) with HPV-16 E6 expression (e.g. SNHG15). This is not surprising, as lncRNAs are known to typically be tissue specific(26). We also found differences in the expression of some lncRNAs between pre-malignant and cancerous cervical cells CaSki and W12/201402 to HCK (e.g. miR205HG). One of the up-regulated lncRNAs in our RNA-seq dataset, FAM83H-AS1 was intriguing. Its expression recently was found to correlate with poor overall survival in a variety of human cancers(27-32) but had not previously been shown in the context of cervical cancer. In addition, it was recently shown to be involved in regulating cellular processes associated with the hallmarks of cancer (e.g. proliferation and migration)(28, 31, 32). To ensure that expression of FAM83H-AS1 is comparable to lncRNAs with well-known functions, we utilized our RNAseq dataset to compare their RPKM values (Table S2).

B. FAM83H-AS1 expression is higher in HPV-16+ pre-malignant and cancerous samples.

Because many lncRNAs are known to be tissue specific(26) and our RNA-seq screening was performed with foreskin keratinocytes expressing HPV-16 E6 oncogene, it was critical to confirm the expression changes of FAM83H-AS1 in epithelial keratinocytes of the cervix where HPV naturally infects(11). Additionally, we considered that the E6 and E7 oncogenes can be synergistic with each other(33), so we needed to develop a model that closely mimics HPV infection by using the entire HPV-16 genome. Therefore, primary human cervical keratinocytes (HCK) were transfected with the entire HPV-16 genome (by releasing the viral genome from a plasmid construct and circularizing it by ligation before transfection), then passaged several times (around 10-15 divisions) for growth selection of HPV positive immortalized cells. HPV-16 E6 and E7 expression was confirmed by RT-PCR (Fig. S2A), as well as p53 degradation through the HPV-16 E6 pathway by Western Blot (Fig. S2B). After confirmation of HPV-16 oncogene expression in these cells, we named them JAMM-16. It was confirmed by qRT-PCR that FAM83H-AS1 is also up-regulated in JAMM-16 cervical keratinocytes expressing the entire HPV-16 genome in comparison to the parental cervical keratinocytes (Fig. 2A). We then used HPV-16 positive low-grade cervical (W12/20863 [episomal HPV-16], W12/201402 [integrated HPV-16]) and carcinoma (CaSki [integrated HPV-16]) cell lines to investigate the expression of FAM83H-AS1. As shown in Figure 2B, we found higher expression of FAM83H-AS1 in all the HPV-16 positive cell lines in comparison to HCK cells. Interestingly, FAM83H-AS1 was expressed at lower levels in HPV-negative

cervical cancer C33A cells compared to HCK as well as HPV-16 positive cervical cells (Fig. 2B). Furthermore, HPV-16 positive and HPV-negative head and neck squamous cell carcinoma (HNSCC) cell lines were compared, and higher expression of FAM83H-AS1 was observed in the HPV-16 positive versus the HPV-negative HNSCC (Fig. 2C). Overall, we conclude that presence of HPV-16 correlates with elevated levels of FAM83H-AS1 expression in early stages of cervical carcinogenesis (newly immortalized JAMM-16 cells and cervical low-grade pre-malignant cell lines) as well as cervical cancer and HNSCC cell lines.

C. FAM83H-AS1 expression is regulated by HPV-16 E6 in a p53-independent, p300-dependent manner.

As shown in Figure 1, foreskin keratinocytes (HEKa) expressing HPV-16 E6 up-regulated FAM83H-AS1 expression. To further confirm HPV-16 E6 regulation of FAM83H-AS1 in cervical cells, cervical keratinocytes (HCK) stably expressing HPV-16 E6 were developed (Fig. S3). To measure if FAM83H-AS1 regulation could be affected by HPV-16 E7 oncogene in synergistic or antagonistic manner, we generated HCK stable cell lines expressing HPV-16 E7 or co-expressing HPV-16 E6 and E7 (Fig. S3). FAM83H-AS1 was up-regulated when cells expressed HPV-16 E6 or HPV-16 E6 and E7, but not when they expressed HPV-16E7 alone (Fig. 3A) suggesting a specific regulation by E6. To further confirm these findings, HPV-16 E6 was knocked down in CaSki (Fig. 3B, Fig. S4A) and W12/201402 (Fig. 3C, Fig. S4B) cell lines by two different siRNAs against HPV-16 E6-E7. FAM83H-AS1 expression was down-regulated after the reduction of HPV-16 E6 expression confirming the regulation of this lncRNA by HPV-16 E6. It is well known that one of the major HPV E6 targets is the tumor suppressor p53(34), which is involved in the regulation of cell proliferation, DNA repair, and apoptosis(35). Interestingly, when we knocked down p53 in HCK by using two different siRNAs, we observed that FAM83H-AS1 expression did not change (Fig. 3D, Fig. S4C), suggesting a regulation by HPV-16 E6 in a p53-independent manner. It is also known that HPV-16 E6 is able to regulate the expression of other important genes in carcinogenesis such as hTERT through the regulation of transcriptional coactivators such as p300(36). By using the UCSC Genome Browser (<https://genome.ucsc.edu/>) to investigate potential p300 binding site in the promoter region of FAM83H-AS1, we found three predicted p300 binding sites (Fig. 3E). In order to measure the potential regulation of FAM83H-AS1 by p300, we used two siRNAs against p300 and quantified FAM83H-AS1 expression. When p300 was knocked down in primary cervical keratinocytes, FAM83H-AS1 expression was also reduced (Fig. 3F, Fig. S4D) suggesting direct and/or indirect p300 regulation of FAM83H-AS1. Previous publications have shown a greater affinity of HPV-16 E6 to interact with p300 in comparison with other high-risk HPV E6 such as HPV-18 E6(37). Interestingly, FAM83H-AS1 expression was found to be up-regulated in HPV-16

positive cell lines (Fig. 2), but down-regulated in HPV-18 positive HeLa cells and HPV-31b positive CIN-612 cells (Fig. S4E). Altogether, we found that FAM83H-AS1 is regulated by HPV-16 E6 independently of p53 but influenced by the presence of p300.

D. FAM83H-AS1 is localized to the nucleus and does not regulate transcription of nearby FAM83H.

The cellular localization of lncRNAs can provide information on their potential function. Nuclear lncRNAs can regulate at the transcriptional level by interacting with critical epigenetic regulators and enhancing chromatin looping, as well as interact with splicing factors to regulate splicing(38). Meanwhile cytoplasmic lncRNAs have been found to function at the post-transcriptional level as competing endogenous RNAs by acting as microRNA sponges and binding to mRNAs leading to the recruitment of RNA binding proteins that promote decay, suppress translation, or factors that initiate translation(39). For this reason, we investigated the cellular localization of FAM83H-AS1 in two HPV-16 positive cervical cell lines by cellular fractionation. We used U6 as a nuclear RNA control and β -actin mature mRNA as a cytoplasmic RNA control. We found significantly higher amplification of FAM83H-AS1 in the nuclear fractions in comparison to the cytoplasmic fractions by qRT-PCR in CaSki (Fig. 4A) and W12/201402 (Fig. 4B) cells, suggesting that FAM83H-AS1 is a nuclear lncRNA. Because many nuclear lncRNAs can act in cis(38), we hypothesized that FAM83H-AS1 could regulate its nearby protein coding gene FAM83H. The lncRNA FAM83H-AS1 and protein coding gene FAM83H share a promoter region but are transcribed in opposite directions(27) (Fig. 3E, Fig. S5). FAM83H is required for the organization of the keratin cytoskeleton in epithelial cells(40) and has been shown over-expressed in different tumor samples compared to their matching normal tissues(41). Interestingly, we found increased expression of FAM83H in HCK expressing HPV-16 E6 in comparison to parental HCK cells (Fig. S5B), but when we transfected an siRNA against FAM83H-AS1 in HPV-16 positive CaSki cells, we were unable to detect changes in FAM83H expression (Fig. S5C), suggesting that FAM83H-AS1 is not involved in regulation of FAM83H expression.

E. FAM83H-AS1 knockdown in cervical cancer cells causes reduced cellular proliferation and migration, as well as induction of apoptosis.

To understand the significance of FAM83H-AS1 in cervical cancer cells, we analyzed the effects on cellular proliferation when FAM83H-AS1 was knocked down by siRNA in CaSki and W12/201402 cervical cells. First, knockdown efficiency of a pool of 4 individual siRNAs (SMARTpool), as well as each of the individual siRNAs, was evaluated by qRT-PCR. Knockdown of FAM83H-AS1 in CaSki with

the siRNA SMARTpool was maintained over a time-course from 24 to 120 hours, which was sufficient for all functional assays conducted (Fig. S6). All of the individual siRNAs, as well as the SMARTpool showed greater than 51% knockdown of FAM83H-AS1 in CaSki (Fig. 5A) and greater than 49% in W12/201402 cells (Fig. S7A). Therefore, we decided to randomly choose two of the individual siRNAs and SMARTpool to knockdown FAM83H-AS1 and monitor cellular functional changes. Two siRNAs against FAM83H-AS1 were transfected independently into CaSki and W12/201402 cells, cultured for 48 hours, and replated to measure cell proliferation by cell counting. In both CaSki and W12/201402 cells, we observed a decrease ($\geq 48\%$) in cell number with knockdown of FAM83H-AS1 compared to control (Fig. 5B, Fig. S7B). Cellular proliferation assay (CCK-8) showed that there was a decrease in cellular proliferation when FAM83H-AS1 is knocked down in CaSki cells and W12/201402 as monitored from 48 hours to 96 hours after replating. We found significantly less proliferation in CaSki (64% decrease) and W12/201402 (73% decrease), in the siRNA FAM83H-AS1 knockdown compared to siRNA control cells at the 96-hour time point (Fig. 5C, Fig. S7C). We observed similar functional changes between the two individual siRNAs and the SMARTpool so for future functional assays we only used the siRNA SMARTpool. In order to identify changes in cell cycle that could explain the differences found in cellular proliferation after FAM83H-AS1 knockdown, we performed cell cycle flow cytometry analysis. CaSki and W12/201402 cells had a significant reduction (43% and 56%, respectively) of cells in S-phase when FAM83H-AS1 was knocked down in comparison to control suggesting that FAM83H-AS1 is important in the G2/S-phase transition (Fig. 5D, Fig. S7D). Other important hallmarks of cancer such as migration and resistance to apoptosis were measured after FAM83H-AS1 knockdown. We found that cellular migration was significantly decreased in CaSki and W12/201402 after siRNA knockdown of FAM83H-AS1 compared to siRNA control (Fig. 5E, Fig. S7E). Also, we measured a significant increase in early and late apoptosis in CaSki and W12/201402 cells with knockdown of FAM83H-AS1 compared to control cells by using Annexin V/PI staining and flow cytometry (Fig. 5F, Fig. S7F and S7G-H). W12/201402 cells with knockdown of FAM83H-AS1 showed a significant increase in necrosis (Fig. S7F), while only a slight increase in necrosis was observed in CaSki cells (Fig. 5F, Fig. S7H). Altogether, we observed significant alterations in cellular proliferation, cell cycle, migration, and apoptosis by the absence of FAM83H-AS1, suggesting an important role in cervical carcinogenesis.

F. FAM83H-AS1 expression is increased in cervical cancer human tissues & is associated with worse overall survival.

In order to extrapolate our findings into a more clinically relevant setting, we extracted RNA from pre-malignant and cervical cancer patient samples and analyzed the expression of FAM83H-AS1 by

qRT-PCR. We found high expression of FAM83H-AS1 in the pre-malignant sample (CIN3) as well as the cervical cancer (CaCx) samples in comparison to normal cervix tissue (Fig. 6A). These findings corroborate our *in vitro* data suggesting an importance of FAM83H-AS1 in clinical tumor samples at different stages of carcinogenesis. We also took advantage of the cervical cancer samples deposit in the TCGA database to compare the expression of FAM83H-AS1 between normal cervix and cervical cancer samples obtained from different cancer stages (120 Stage I, 35 Stage II, 30 Stage III, 7 Stage IV, 4 Stage unavailable). The TCGA data showed elevated expression of FAM83H-AS1 (RPKM values) in cervical cancer patients compared to normal cervix control (Fig. 6B). This coincides with our previous observations of FAM83H-AS1 expression being higher in cervical cancer cells lines (Fig. 2B). Finally, we used the TCGA data set from the TANRIC database to divide cervical cancer patients into high versus low expression groups and measured overall survival based on FAM83H-AS1 expression. Interestingly, we found that patients with higher FAM83H-AS1 expression yielded a worse overall survival than patients with lower FAM83H-AS1 expression suggesting a biological importance of this lncRNA in cervical cancer reflected in patients' clinical outcomes (Fig. 6C). Overall, we conclude that FAM83H-AS1 expression is elevated in cervical cancer patients and high expression correlates with overall poor survival.

IV. Discussion

4.5% of all cancers worldwide are attributable to HPV infection. Almost all cervical cancers and a substantial amount of other anogenital and oropharyngeal cancers have been found to be infected by high-risk HPVs. HPV-16 and -18 contribute to 73% of HPV-associated cancers(42), implying a higher ability to induce tumorigenesis in comparison to other types of HPVs. Although the prevalence of HPV-associated cancers has decreased due to development of the preventative vaccine and early detection screening methods(2), there is still a great need for prognostic and therapeutic options specially to people already infected with HPV as well as those affected in less developed countries where the access to the vaccine is limited.

Long noncoding RNAs (lncRNAs) have been shown to regulate a variety of critical cellular processes, including transcription and chromosome remodeling(12-14). Dysregulation of lncRNAs has been shown to be associated with the development and progression of many cancers(15-17), and interestingly they are commonly tissue specific(26) and only altered in one cancer type(32). Therefore, lncRNAs are currently being studied in the context of biomarkers for diagnosis and prognosis of cancer, as well as therapeutic targets(15-18).

Previous studies have shown that the high-risk HPV E6 protein expressed is clearly involved in the progression of carcinogenesis. HPV E6-regulation of non-coding RNAs such as microRNAs has been well studied(43), however long non-coding RNA regulation by high-risk HPV E6 needs to be studied further; there are only a couple lncRNAs shown to be specifically HPV-16 E6 regulated, including MALAT1 and CCEPR(22, 23). To add to this field of study, we sought out to identify an HPV-16 E6 regulated gene that was altered from the early stages of HPV infection until carcinogenesis and therefore we considered to be important in both the development and progression of carcinogenesis. Thus, we developed a new HPV-16 positive cell line referred to as JAMM-16 to represent early infection, but also analyzed expression in established HPV-16 positive cervical cell lines such as CaSki and W12 cells as well as pre-malignant and malignant cervical tumor samples. We found FAM83H-AS1 overexpression in W12/20863 and W12/201402 (which came from a CIN2 tumor) similar to CaSki cells (cervical carcinoma). Additionally, in Figure 6A, the CIN3 (considered Stage 0 cervical carcinoma) patient sample shows similar expression to later stage cervical carcinoma (CaCx) samples. For this reason, we believe that FAM83H-AS1 up-regulation in pre-malignant cervical samples could be linked to the expression of viral oncogenes in early HPV infection.

From the host lncRNAs altered by HPV-16 E6 in our RNA-seq analysis (Fig. 1), we found several of these lncRNAs previously described to be altered in cervical cancers, confirming that our data aligned with former studies. For example, it was previously found that decreased expression of GAS5 is associated with poor prognosis of cervical cancer patients(44, 45) as well as is tumor suppressive in other types of cancer such as breast cancer(46, 47) and prostate cancer(47). Furthermore, GAS5 expression was also found altered *in vitro* in HPV-16 positive CaSki cells(44) and HPV-18 positive HeLa cells(45). Another lncRNAs affected in our study was H19. It was shown previously that DNA methylation alterations at the IGF2/H19 imprinted domain may mediate the association between HPV and invasive cervical cancer(48) and high H19 expression has also been shown to be predictive of poor prognosis in cervical cancer(49) as well as in a variety of other human cancers, including HNSCC(50) and breast cancer(51).

A class of lncRNAs known as onco-lncRNAs were also interesting to us because, as the majority of lncRNAs are tissue specific(26), this group of lncRNAs exhibit differential expression across multiple cancers and are hypothesized to have conserved oncogenic or tumor suppressive functions. One such onco-lncRNA which was found in our RNA sequencing analysis of an E6-regulated lncRNA is onco-lncRNA-3, referred to as FAM83H-AS1. This lncRNA is transcribed from chromosome 8 and its function

in normal cells is unknown. It was first characterized in 2015(32), and as of now multiple publications have shown increased expression of FAM83H-AS1 in breast(29, 32), lung(31, 32), colorectal(28, 30, 32), kidney(32), bladder(32), and pancreatic cancers(27) and increased expression correlates with worse overall survival in most of these cancers(27, 29-31). Our findings show for the first time that FAM83H-AS1 is overexpressed in human cervical cancer (CESC) tissues and high expression in patients correlates with poor overall survival (Fig. 6).

According to previous studies, FAM83H-AS1 is an epithelial lncRNA(27) supporting our data obtained from foreskin and cervical keratinocytes. Determining the localization of a lncRNA can predict functionality of the lncRNA; our findings that FAM83H-AS1 is localized in the nucleus of cervical cancer cells (Fig. 4) is consistent with previous findings of its nuclear localization in lung cancer cells¹⁴. Functionally, it has been found to be co-expressed with protein coding genes that were enriched for cell cycle-related genes(32), and knockdown of FAM83H-AS1 altered cell cycle(31, 32), proliferation(28, 31), migration(28, 31), invasion(31), and apoptosis(28) in certain cancers. Our group shows here that in the context of cervical cancer, FAM83H-AS1 is involved in cell cycle, proliferation, migration, and apoptosis (Fig. 5). It is unknown if FAM83H-AS1 elicits its functions in cis or trans, but our findings suggest that FAM83H-AS1 does not elicit cis regulation on the nearby protein coding gene FAM83H (Fig. S5), which is up-regulated in a variety of human cancers. For this reason, it will be interesting in future studies to identify the protein, RNA, and/or DNA interactions of FAM83H-AS1 in cervical cancers.

Previously, it was shown that FAM83H-AS1 regulates MET/EGFR signaling in lung cancer cells(31) and that when FAM83H-AS1 was downregulated it exhibited an anti-proliferative role by suppressing the Notch signaling pathway in colorectal cancer(28). To elucidate additional downstream targets of FAM83H-AS1, a group recently conducted RNA-seq on a pancreatic cancer cell line with siRNA knockdown of FAM83H-AS1 compared to control and identified gene alterations (78 activated and 68 inhibited targets)(27). Our group plans to determine if these downstream regulators are also involved in FAM83H-AS1 mediated functional changes observed in cervical cancer cells.

HPV-16 E6 and -18 E6 are well known to contribute to the degradation of p53, however, it is important to note that HPV-16 and HPV-18 vary in their interactions with other proteins to regulate carcinogenesis. For example, previous studies have shown that HPV-16 E6 directly interacts with CBP/p300(8, 10, 37), but HPV-18 E6 appears to be unable to interact with p300(37). This could be a possible explanation for variation in FAM83H-AS1 expression between HPV-16 and -18 positive cancers

observed in our study (Fig. 2, Fig. S4E). Supporting our data, another publication previously showed low expression of FAM83H-AS1 in HeLa (HPV-18 positive) cells(27). This information led us to elucidate the mechanism of FAM83H-AS1 up-regulation by E6 in a p53-independent and p300-dependent manner (Fig. 3, Fig. S4). This regulation is interesting because the majority of E6 regulation of several coding and non-coding genes is primarily through the p53 pathway. Interestingly, a previous study showed that overexpression of cyclooxygenase (COX-2) gene was the result of the recruitment of p300 to its promoter region via the overexpression of HPV-16 E6 protein in CaSki cells as well as the exogenous expression of HPV-16 E6 in HPV-negative C-33A cells(52). Recently, p300 inhibitors such as C646 have been shown to be good candidates as anti-cervical cancer drugs, demonstrating the importance of p300 not only in the regulation of host genes but also of HPV viral genes(53).

In summary, the identification of FAM83H-AS1 up-regulation in the early stages of cervical carcinogenesis, correlation with overall survival in cervical cancer, and involvement in different hallmarks of cancer contributes further evidence of the importance of this lncRNA in cancer. Further studies on this lncRNA could enhance the use of FAM83H-AS1 as a potential biomarker or therapeutic target in multiple cancers.

V. Methods

Detailed experimental protocols are described in the Supplementary Methods section. All experiments were performed in compliance with the Institutional Biosafety Committee at West Virginia University, number 15-03-03.

A. Cells

The following cell lines were used: human primary foreskin keratinocytes (HEKa) (Invitrogen, C-005-5C); human primary cervical keratinocytes (HCK) and J2-3T3 murine fibroblast feeder cells (obtained from Dr. Alison McBride's laboratory, NIH, Bethesda, MD); 3T3M murine fibroblast feeder cells, as well as CaSki (HPV-16 positive), HeLa (HPV-18 positive), and C-33A (HPV negative) cervical carcinoma cells (obtained from Dr. Daniel DiMaio's laboratory, Yale University, New Haven, CT); W12/201402 (HPV-16 positive) and W12/20863 (HPV-16 positive) pre-malignant cervical cells (obtained from Paul F. Lambert, University of Wisconsin-Madison, Madison, WI); CIN-612 (HPV-31b positive) (obtained from Dr. Laimonis A. Laimins' laboratory, Northwestern University, Chicago, IL); UMSCC-1 (HPV negative), UMSCC-47 (HPV-16 positive), and UMSCC-104 (HPV-16 positive) head

and neck squamous cell carcinoma (HNSCC) cell lines (obtained from Dr. Scott A. Weed's laboratory, West Virginia University, Morgantown, WV). Further details are in the Supplementary Methods section.

B. High-throughput RNA sequencing

Three replicates each of human foreskin keratinocytes (HEKa) stably expressing HPV-16 E6 or GFP were sent for RNA high-throughput sequencing (Illumina). FASTQ files were subsequently imported into Strand NGS suites for analysis. Reads were aligned to the human hg19 reference genomes using the Bowtie algorithm. These were then quantified against Ensemble transcript and including small and lncRNA annotations. Any lncRNA that were detected in human subjects were used for further analysis. Raw lncRNA counts were then normalized to the total number of lncRNA reads per sample and expression values calculated against the control samples. Further Mapping rate visualization done using Strand NGS software. In order to determine the importance of the lncRNAs obtained from the RNA-seq analysis, we used the following filtering strategy: First, only lncRNAs with reasonable expression (RPKM greater than 1) were analyzed further. Then, we used The Atlas of Noncoding RNAs in Cancer (TANRIC, MD Anderson Cancer Center)(54), which contains 297 sequenced human cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) patient data from The Cancer Genome Atlas (TCGA), to analyze the expression and clinical outcomes of these lncRNAs in patient samples. Finally, we searched previous publications to identify lncRNAs altered in other types of cancer and/or involved in hallmarks of cancer. To increase our novelty, we eliminated lncRNAs that were previously shown to be involved specifically in cervical cancer (e.g. H19).

C. Functional Analysis

For cell counting experiments, CaSki cells were transiently transfected with Lincode Human FAM83H-AS1 siRNA SMARTpool (Dharmacon, R-188909-00-000), Individual: Lincode FAM83H-AS1 siRNA (N-188909-02-0002 and N-188909-04-0002), or Lincode Non-targeting siRNA #1 (Dharmacon, D-001320-01-05) using Lipofectamine® RNAiMAX according to manufacturer's instructions (Invitrogen).

CaSki cells were cultured with siRNA-containing media for 48 hours, re-plated in equal cell numbers (200,000 cells/well of 6-well), cultured for another 48 hours, and attached cells were re-counted with a hemacytometer.

For all other functional assays (CCK-8 cell proliferation, FACS cell cycle, transwell migration, annexin V-FITC/PI apoptosis), CaSki cells were transiently transfected with Lincode Human FAM83H-AS1 siRNA SMARTpool or Lincode Non-targeting siRNA #1 (Dharmacon, D-001320-01-05) using Lipofectamine® RNAiMAX according to manufacturer's instructions (Invitrogen). Cells were incubated with siRNA-containing media for 24 hours then re-plated in equal cell numbers for to initiate experiments described below. To monitor cell proliferation, transfected cells were plated in 96-well plates and after 24, 48, 72, and 96 hours analyzed with CCK-8 kit (Sigma-Aldrich) according to manufacturer's protocol. Alterations in cell cycle were determined by flow cytometry propidium iodide DNA staining. Transfected cells were plated in media containing 10% fetal bovine serum (FBS). Cells were allowed to attach and then were serum-starved for 24 hours. Samples were then fixed with ethanol, stained with propidium iodide, and analyzed by flow cytometry (Fortessa S10). For transwell migration assay, transfected cells were seeded onto upper chambers of transwell inserts (8µm pore size) with 20% FBS chemoattractant in the lower chamber of 24-well plate. After 48 hours, migrated cells located on the underside of the transwell insert were stained with 0.5% crystal violet in methanol. Migrated cells were quantified using ImageJ software. To monitor apoptosis, transfected cells were plated and incubated 24, 48, and 72 hours. At desired time point, attached and floating cells were pelleted and co-stained with annexin V-FITC and propidium iodine and immediately analyzed by flow cytometry (Fortessa S10).

VI. Data Availability

The RNA-seq raw data generated during and/or analyzed during the current study are available in the Gene Expression Omnibus (GEO-NCBI) repository, accession number: GSE115334.

VII. References

1. **Cubie HA.** 2013. Diseases associated with human papillomavirus infection. *Virology* **445**:21-34.
2. **Moody CA, Laimins LA.** 2010. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer* **10**:550-560.
3. **Yim EK, Park JS.** 2005. The role of HPV E6 and E7 oncoproteins in HPV-associated cervical carcinogenesis. *Cancer Res Treat* **37**:319-324.
4. **Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM.** 1990. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* **63**:1129-1136.
5. **Boyer SN, Wazer DE, Band V.** 1996. E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Res* **56**:4620-4624.
6. **Roman A, Munger K.** 2013. The papillomavirus E7 proteins. *Virology* **445**:138-168.

7. **Vande Pol SB, Klingelhutz AJ.** 2013. Papillomavirus E6 oncoproteins. *Virology* **445**:115-137.
8. **Zimmermann H, Degenkolbe R, Bernard HU, O'Connor MJ.** 1999. The human papillomavirus type 16 E6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator CBP/p300. *J Virol* **73**:6209-6219.
9. **Scheffner M, Huibregtse JM, Vierstra RD, Howley PM.** 1993. The HPV-16 E6 and E6-AP complex functions as a ubiquitin-protein ligase in the ubiquitination of p53. *Cell* **75**:495-505.
10. **Patel D, Huang SM, Baglia LA, McCance DJ.** 1999. The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. *Embo j* **18**:5061-5072.
11. **Moody C.** 2017. Mechanisms by which HPV Induces a Replication Competent Environment in Differentiating Keratinocytes. *Viruses* **9**.
12. **Long Y, Wang X, Youmans DT, Cech TR.** 2017. How do lncRNAs regulate transcription? *Sci Adv* **3**:eaao2110.
13. **Marchese FP, Raimondi I, Huarte M.** 2017. The multidimensional mechanisms of long noncoding RNA function. *Genome Biol* **18**:206.
14. **Wang KC, Chang HY.** 2011. Molecular mechanisms of long noncoding RNAs. *Mol Cell* **43**:904-914.
15. **Chen L, Dzakah EE, Shan G.** 2018. Targetable long non-coding RNAs in cancer treatments. *Cancer Lett* **418**:119-124.
16. **Vitiello M, Tuccoli A, Poliseno L.** 2015. Long non-coding RNAs in cancer: implications for personalized therapy. *Cell Oncol (Dordr)* **38**:17-28.
17. **Gutschner T, Diederichs S.** 2012. The hallmarks of cancer: a long non-coding RNA point of view. *RNA Biol* **9**:703-719.
18. **Arun G, Diermeier SD, Spector DL.** 2018. Therapeutic Targeting of Long Non-Coding RNAs in Cancer. *Trends Mol Med* **24**:257-277.
19. **Dong J, Su M, Chang W, Zhang K, Wu S, Xu T.** 2017. Long non-coding RNAs on the stage of cervical cancer (Review). *Oncol Rep* **38**:1923-1931.
20. **Song J, Ye A, Jiang E, Yin X, Chen Z, Bai G, Zhou Y, Liu J.** 2018. Reconstruction and analysis of the aberrant lncRNA-miRNA-mRNA network based on competitive endogenous RNA in CESC. *J Cell Biochem* doi:10.1002/jcb.26850.
21. **Zhu H, Chen X, Hu Y, Shi Z, Zhou Q, Zheng J, Wang Y.** 2017. Long non-coding RNA expression profile in cervical cancer tissues. *Oncol Lett* **14**:1379-1386.
22. **Sharma S, Munger K.** 2018. Expression of the cervical carcinoma expressed PCNA regulatory (CCEPR) long noncoding RNA is driven by the human papillomavirus E6 protein and modulates cell proliferation independent of PCNA. *Virology* **518**:8-13.
23. **Liu S, Song L, Yao H, Zhang L, Xu D, Gao F, Li Q.** 2016. MiR-375 Is Epigenetically Downregulated by HPV-16 E6 Mediated DNMT1 Upregulation and Modulates EMT of Cervical Cancer Cells by Suppressing lncRNA MALAT1. *PLoS One* **11**:e0163460.
24. **Gibb EA, Becker-Santos DD, Enfield KS, Guillaud M, Niekerk D, Matisic JP, Macaulay CE, Lam WL.** 2012. Aberrant expression of long noncoding RNAs in cervical intraepithelial neoplasia. *Int J Gynecol Cancer* **22**:1557-1563.
25. **Iancu IV, Anton G, Botezatu A, Huica I, Nastase A, Socolov DG, Stanescu AD, Dima SO, Bacalbasa N, Plesa A.** 2017. LINC01101 and LINC00277 expression levels as novel factors in HPV-induced cervical neoplasia. *J Cell Mol Med* **21**:3787-3794.
26. **Yan X, Hu Z, Feng Y, Hu X, Yuan J, Zhao SD, Zhang Y, Yang L, Shan W, He Q, Fan L, Kandalaft LE, Tanyi JL, Li C, Yuan CX, Zhang D, Yuan H, Hua K, Lu Y, Katsaros D, Huang Q, Montone K, Fan Y, Coukos G, Boyd J, Sood AK, Rebbeck T, Mills GB, Dang CV, Zhang L.** 2015. Comprehensive Genomic Characterization of Long Non-coding RNAs across Human Cancers. *Cancer Cell* **28**:529-540.
27. **Arnes L, Liu Z, Wang J, Carlo Maurer H, Sagalovskiy I, Sanchez-Martin M, Bommakanti N, Garofalo DC, Balderes DA, Sussel L, Olive KP, Rabadan R.** 2018. Comprehensive

- characterisation of compartment-specific long non-coding RNAs associated with pancreatic ductal adenocarcinoma. *Gut* doi:10.1136/gutjnl-2017-314353.
28. **Lu S, Dong W, Zhao P, Liu Z.** 2018. lncRNA FAM83H-AS1 is associated with the prognosis of colorectal carcinoma and promotes cell proliferation by targeting the Notch signaling pathway. *Oncol Lett* **15**:1861-1868.
 29. **Yang F, Lv SX, Lv L, Liu YH, Dong SY, Yao ZH, Dai XX, Zhang XH, Wang OC.** 2016. Identification of lncRNA FAM83H-AS1 as a novel prognostic marker in luminal subtype breast cancer. *Onco Targets Ther* **9**:7039-7045.
 30. **Yang L, Xu L, Wang Q, Wang M, An G.** 2016. Dysregulation of long non-coding RNA profiles in human colorectal cancer and its association with overall survival. *Oncol Lett* **12**:4068-4074.
 31. **Zhang J, Feng S, Su W, Bai S, Xiao L, Wang L, Thomas DG, Lin J, Reddy RM, Carrott PW, Lynch WR, Chang AC, Beer DG, Guo YM, Chen G.** 2017. Overexpression of FAM83H-AS1 indicates poor patient survival and knockdown impairs cell proliferation and invasion via MET/EGFR signaling in lung cancer. *Sci Rep* **7**:42819.
 32. **Cabanski CR, White NM, Dang HX, Silva-Fisher JM, Rauck CE, Cicka D, Maher CA.** 2015. Pan-cancer transcriptome analysis reveals long noncoding RNAs with conserved function. *RNA Biol* **12**:628-642.
 33. **Jabbar S, Strati K, Shin MK, Pitot HC, Lambert PF.** 2010. Human papillomavirus type 16 E6 and E7 oncoproteins act synergistically to cause head and neck cancer in mice. *Virology* **407**:60-67.
 34. **Munger K, Howley PM.** 2002. Human papillomavirus immortalization and transformation functions. *Virus Res* **89**:213-228.
 35. **Biegging KT, Mello SS, Attardi LD.** 2014. Unravelling mechanisms of p53-mediated tumour suppression. *Nat Rev Cancer* **14**:359-370.
 36. **James MA, Lee JH, Klingelhutz AJ.** 2006. HPV16-E6 associated hTERT promoter acetylation is E6AP dependent, increased in later passage cells and enhanced by loss of p300. *Int J Cancer* **119**:1878-1885.
 37. **White EA, Kramer RE, Tan MJ, Hayes SD, Harper JW, Howley PM.** 2012. Comprehensive analysis of host cellular interactions with human papillomavirus E6 proteins identifies new E6 binding partners and reflects viral diversity. *J Virol* **86**:13174-13186.
 38. **Yu B, Shan G.** 2016. Functions of long noncoding RNAs in the nucleus. *Nucleus* **7**:155-166.
 39. **Noh JH, Kim KM, McClusky WG, Abdelmohsen K, Gorospe M.** 2018. Cytoplasmic functions of long noncoding RNAs. *Wiley Interdiscip Rev RNA* **9**:e1471.
 40. **Kuga T, Sasaki M, Mikami T, Miake Y, Adachi J, Shimizu M, Saito Y, Koura M, Takeda Y, Matsuda J, Tomonaga T, Nakayama Y.** 2016. FAM83H and casein kinase I regulate the organization of the keratin cytoskeleton and formation of desmosomes. *Sci Rep* **6**:26557.
 41. **Snijders AM, Lee SY, Hang B, Hao W, Bissell MJ, Mao JH.** 2017. FAM83 family oncogenes are broadly involved in human cancers: an integrative multi-omics approach. *Mol Oncol* **11**:167-179.
 42. **de Martel C, Plummer M, Vignat J, Franceschi S.** 2017. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int J Cancer* **141**:664-670.
 43. **Martinez I, Gardiner AS, Board KF, Monzon FA, Edwards RP, Khan SA.** 2008. Human papillomavirus type 16 reduces the expression of microRNA-218 in cervical carcinoma cells. *Oncogene* **27**:2575-2582.
 44. **Cao S, Liu W, Li F, Zhao W, Qin C.** 2014. Decreased expression of lncRNA GAS5 predicts a poor prognosis in cervical cancer. *Int J Clin Exp Pathol* **7**:6776-6783.
 45. **Li Y, Wan YP, Bai Y.** 2018. Correlation between long strand non-coding RNA GAS5 expression and prognosis of cervical cancer patients. *Eur Rev Med Pharmacol Sci* **22**:943-949.
 46. **Mourtada-Maarabouni M, Pickard MR, Hedge VL, Farzaneh F, Williams GT.** 2009. GAS5, a non-protein-coding RNA, controls apoptosis and is downregulated in breast cancer. *Oncogene* **28**:195-208.

47. **Ma C, Shi X, Zhu Q, Li Q, Liu Y, Yao Y, Song Y.** 2016. The growth arrest-specific transcript 5 (GAS5): a pivotal tumor suppressor long noncoding RNA in human cancers. *Tumour Biol* **37**:1437-1444.
48. **Vidal AC, Henry NM, Murphy SK, Oneko O, Nye M, Bartlett JA, Overcash F, Huang Z, Wang F, Mlay P, Obure J, Smith J, Vasquez B, Swai B, Hernandez B, Hoyo C.** 2014. PEG1/MEST and IGF2 DNA methylation in CIN and in cervical cancer. *Clin Transl Oncol* **16**:266-272.
49. **Peng L, Yuan XQ, Liu ZY, Li WL, Zhang CY, Zhang YQ, Pan X, Chen J, Li YH, Li GC.** 2017. High lncRNA H19 expression as prognostic indicator: data mining in female cancers and polling analysis in non-female cancers. *Oncotarget* **8**:1655-1667.
50. **Guan GF, Zhang DJ, Wen LJ, Xin D, Liu Y, Yu DJ, Su K, Zhu L, Guo YY, Wang K.** 2016. Overexpression of lncRNA H19/miR-675 promotes tumorigenesis in head and neck squamous cell carcinoma. *Int J Med Sci* **13**:914-922.
51. **Shima H, Kida K, Adachi S, Yamada A, Sugae S, Narui K, Miyagi Y, Nishi M, Ryo A, Murata S, Taniguchi H, Ichikawa Y, Ishikawa T, Endo I.** 2018. Lnc RNA H19 is associated with poor prognosis in breast cancer patients and promotes cancer stemness. *Breast Cancer Res Treat* doi:10.1007/s10549-018-4793-z.
52. **Subbaramaiah K, Dannenberg AJ.** 2007. Cyclooxygenase-2 transcription is regulated by human papillomavirus 16 E6 and E7 oncoproteins: evidence of a corepressor/coactivator exchange. *Cancer Res* **67**:3976-3985.
53. **He H, Lai Y, Hao Y, Liu Y, Zhang Z, Liu X, Guo C, Zhang M, Zhou H, Wang N, Luo XG, Huo L, Ma W, Zhang TC.** 2017. Selective p300 inhibitor C646 inhibited HPV E6-E7 genes, altered glucose metabolism and induced apoptosis in cervical cancer cells. *Eur J Pharmacol* **812**:206-215.
54. **Li J, Han L, Roebuck P, Diao L, Liu L, Yuan Y, Weinstein JN, Liang H.** 2015. TANRIC: An Interactive Open Platform to Explore the Function of lncRNAs in Cancer. *Cancer Res* **75**:3728-3737.

VIII. Acknowledgement

This work was supported in part by WVU Cancer Institute, National Cancer Institute (NCI) Plan (2V882), American Cancer Society IRG Internal Pilot Funding (IRG-09-061-04), WVCTSI Award (NIH/NIGMS Award Number U54GM104942), Tumor Microenvironment (TME) CoBRE Grant (NIH/NIGMS P20GM121322). Karen Hayes was supported in part by The Ladies Auxiliary to the VFW of the United States Cancer Research Fellowship (CK003229). Neil Infante (WVU Bioinformatics Core) assisted with the high-throughput RNA sequencing data analysis. We acknowledge Molly Lovern, a West Virginia IDeA Network of Biomedical Research Excellence (WV-INBRE) undergraduate student, and WV-INBRE funding source [NIH (NIGMS) Grant #P20GM103434] assisted in developing the JAMM-16 cell line. Sijin Wen (WVU Biostatistics Core) assisted with biostatistics on the TANRIC/TCGA data analysis. We thank Kathy Brundage in the WVU Flow Cytometry Core for assistance with FACS analysis, along with funding sources for the Fortessa S10 (OD016165) and overall core support (MBRCC CoBRE grant: GM103488). We thank Alison McBride at the NIH for kindly providing primary human cervical keratinocytes.

IX. Author Contributions:

J.A.B., K.E.H. and I.M. designed research; S.K. provided initial analysis and collected cervical tissue samples; J.A.B, K.E.H., T.B., A.D.H, and I.M. performed research; J.A.B, K.E.H., T.B., A.D.H, R.J. P.R.L., and I.M. analyzed data; J.A.B. and I.M. wrote the paper.

X. Additional Information:

The authors declare no competing interests.

XI. Figure Legends:

Figure 1: Differential expression of host lncRNAs after expression of HPV-16 E6 in primary foreskin keratinocytes. (A) Waterfall plot of host lncRNAs altered 1.5-fold or greater in primary foreskin keratinocytes (HEKa) expressing HPV-16 E6 compared to uninfected HEKa by high-throughput RNA sequencing (RNA-seq) analysis. Triplicates for each sample were sent for RNA-seq. (B) qRT-PCR validation of representative differentially expressed host lncRNAs found by RNA-seq analysis. Red bars represent the lncRNAs up-regulated with HPV-16 E6 expression, and green bars represent the lncRNAs down-regulated with HPV-16 E6 expression. GAPDH mRNA was used to normalize the qRT-PCR analyses, which are shown relative to uninfected HEKa (grey bars).

Figure 2: Increased FAM83H-AS1 expression in primary cervical keratinocytes containing the HPV-16 genome as well as in HPV-16 positive cervical cancer and head and neck squamous cell carcinoma cell lines. (A) qRT-PCR analysis showing the increase of FAM83H-AS1 expression in newly immortalized cervical keratinocytes expressing entire HPV-16 genome (JAMM-16) compared to uninfected primary cervical keratinocytes (HCK). (B) qRT-PCR analysis showing the increase of FAM83H-AS1 expression in HPV-16 positive cervical cell lines (CaSki, W12/20863, W12/201402) and decrease of FAM83H-AS1 in HPV negative cervical cancer cell line (C-33A) compared to uninfected cervical keratinocytes (HCK). (C) qRT-PCR analysis showing the increase of FAM83H-AS1 expression in HPV-16 positive head and neck squamous cell carcinoma (HNSCC) cell lines (UMSCC-47 and UMSCC-104) compared to HPV negative HNSCC cell line (UMSCC-1). All graphs in the figure show the average of two individual experiments. Similar results were obtained in at least three independent experiments. GAPDH mRNA was used to normalize the qRT-PCR analyses. Two-tailed t test results are indicated as ** $p \leq 0.01$.

Figure 3: Regulation of FAM83H-AS1 expression by HPV-16 E6 in a p53-independent, p300-dependent manner. (A) FAM83H-AS1 expression by qRT-PCR analysis in cervical keratinocytes (HCK) stably individually expressing HPV-16 E6 or E7, or co-expressing E6/E7 compared to GFP control. (B-C) qRT-PCR analysis of HPV-16 E6 and FAM83H-AS1 expression in HPV-16 positive CaSki (B) and W12/201402 (C) cervical cell lines transfected with an siRNA against HPV-16E6 compared to siRNA control. (D) qRT-PCR analysis of p53 and FAM83H-AS1 expression in HCK transfected with an siRNA against p53 compared to siRNA control. (E) Genome representative image showing location of FAM83H, FAM83H-AS1, and three predictive p300 binding sites in FAM83H-AS1 promoter region. (F) p300 and FAM83H-AS1 expression in HCK transfected with an siRNA against p300 compared to siRNA control. All graphs in the figure show the average of two individual experiments. Similar results were obtained in at least three independent experiments. GAPDH mRNA was used to normalize the qRT-PCR analyses. Two-tailed t test results are indicated as * $p \leq 0.05$ and ** $p \leq 0.01$. CTRL, control.

Figure 4: FAM83H-AS1 is localized in the nucleus in cervical pre-malignant and cancerous cell lines. (A) qRT-PCR of FAM83H-AS1 expression in fractionated HPV-16 positive cervical cancer CaSki (A) and pre-malignant W12/201402 (B) cell lines. U6 small nuclear RNA (snRNA) was used as a nuclear control RNA and mature Beta Actin was used as a cytoplasmic control RNA. Representative images; similar results were obtained in at least three independent experiments. Normalization was done using *C. Elegans* total RNA as an exogenous spike for the amplification of worm-specific *ama-1* gene. Two-tailed t test results are indicated as ** $p \leq 0.01$.

Figure 5: FAM83H-AS1 knockdown altered cell proliferation, migration, and apoptosis in CaSki cells. (A) Knockdown efficiency of individual and SMARTpool siRNA against FAM83H-AS1 in HPV-16 positive CaSki cell line, measured by qRT-PCR analysis. Because of variations in the expression of GAPDH after the knockdown of FAM83H-AS1, we used UBC mRNA to normalize the qRT-PCR analyses. The graph shows average of two individual experiments. (B) CaSki cells were transfected with individual siRNAs against FAM83H-AS1, siRNA SMARTpool against FAM83H-AS1, or siRNA control

for 48 hours. Cells were then re-plated in equal numbers (200,000 cells/well, represented by dashed line in graph) and cultured another 48 hours prior to re-counting attached cells. Data were obtained in triplicate, and the graph shows the average of two individual experiments. (C) CaSki cells were transfected with individual siRNAs against FAM83H-AS1, siRNA SMARTpool against FAM83H-AS1, or siRNA control for 24 hours then plated in equal numbers. Transfected cells were analyzed for cellular proliferation assessment by CCK-8 assay at 48, 72, and 96 hours post-plating. The graph shows the average of two individual experiments; similar results were obtained in three independent experiments. (D-F) CaSki cells were transfected with siRNA SMARTpool against FAM83H-AS1 or siRNA control for 24 hours then plated in equal numbers for experiments. (D) Transfected cells were analyzed for cell cycle alterations by FACS analysis. CaSki cells with knockdown of FAM83H-AS1 exhibit less cells in S-phase of cell cycle compared to control cells. The graph shows the average of two individual experiments; similar results were obtained in three independent experiments. (E) Transwell migration of transfected cells was analyzed 48 hours post-plating in upper chamber with chemoattractant in lower chamber. The graph shows the average of three individual experiments. (F) Transfected CaSki cells were collected at 1, 2, and 3 days post-plating, stained with Annexin V/PI, and analyzed by flow cytometry to show alterations in apoptosis compared to siRNA control. The graph shows the average of three individual experiments. Two-tailed t test results are indicated as * $p \leq 0.05$ and ** $p \leq 0.01$.

Figure 6: FAM83H-AS1 expression is increased in human cervical cancer tissues and correlates with poor overall survival. (A) Increased FAM83H-AS1 expression in human cervical cancer and cervical intraepithelial neoplasia (CIN) stage 3 patient samples compared to non-cancerous cervical tissue as measured by qRT-PCR analysis. (B) The Cancer Genome Atlas (TCGA) analysis of FAM83H-AS1 RPKM values in cervical cancer tissues (n=196) compared to non-cancerous tissues (n=3). (C) Survival plot of cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) patient data mined from TCGA with low (n=22) vs. high (n=174) expression of FAM83H-AS1. High expression of FAM83H-AS1 expression correlates with worse overall survival in CESC patients.

Figure 1:

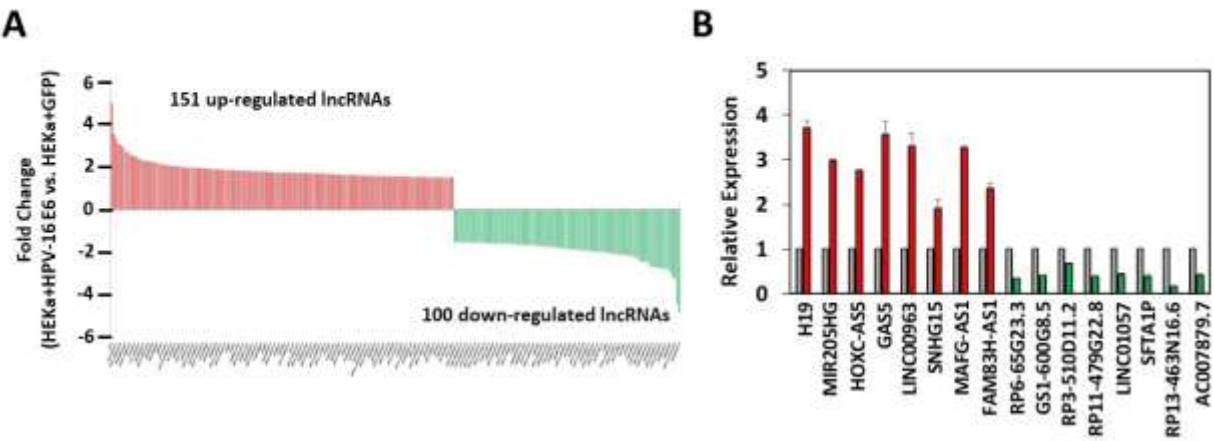


Figure 2:

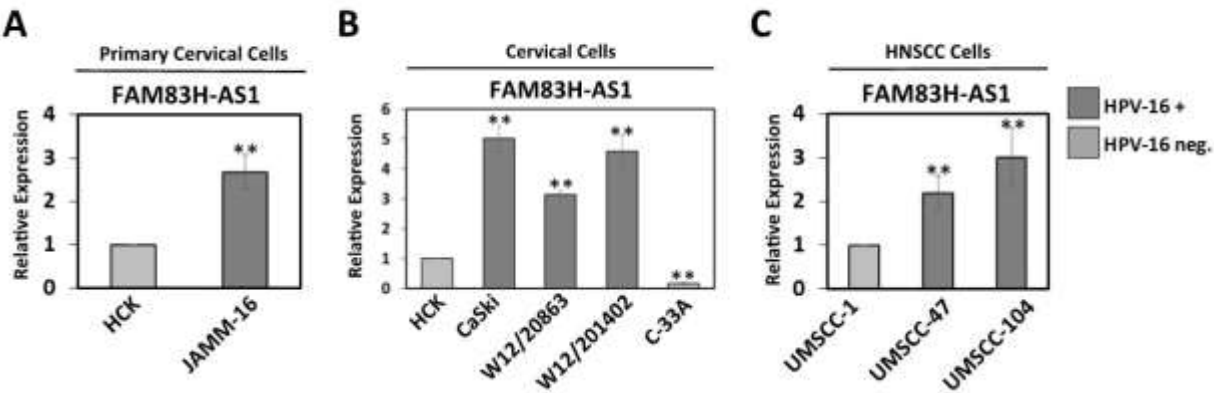


Figure 3:

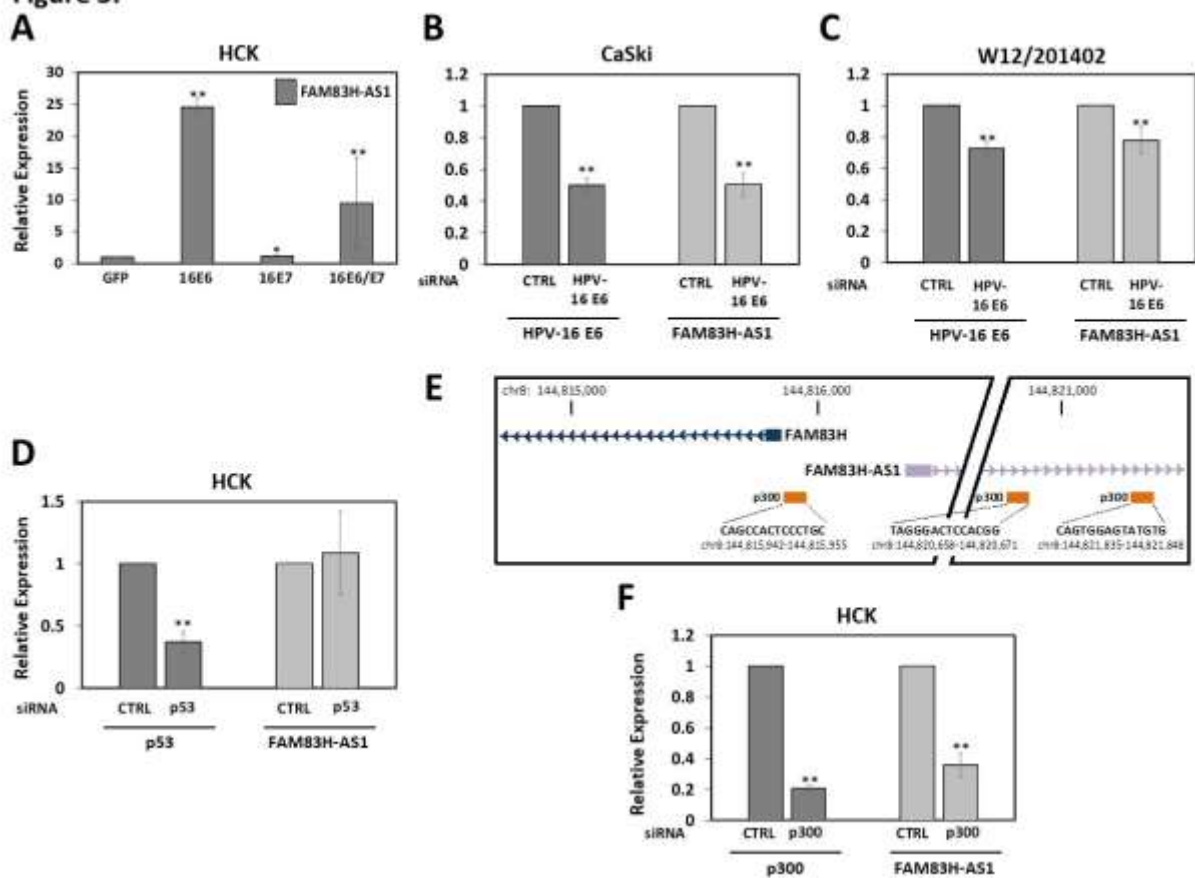


Figure 4:

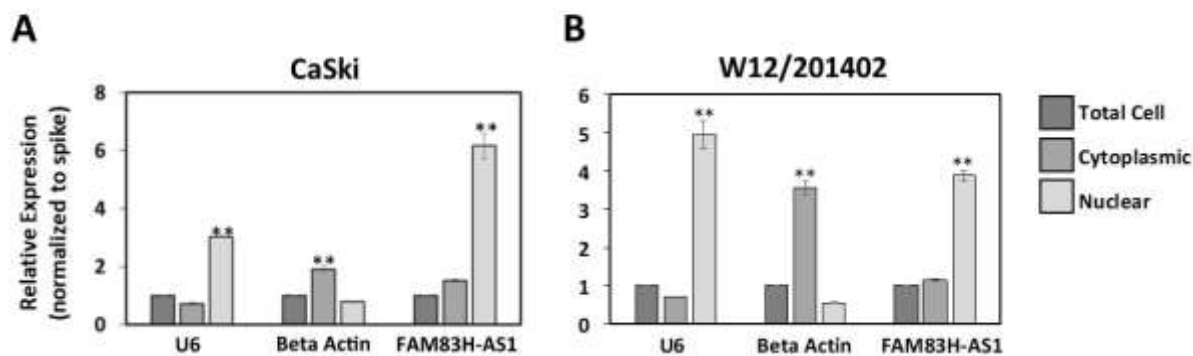


Figure 5:

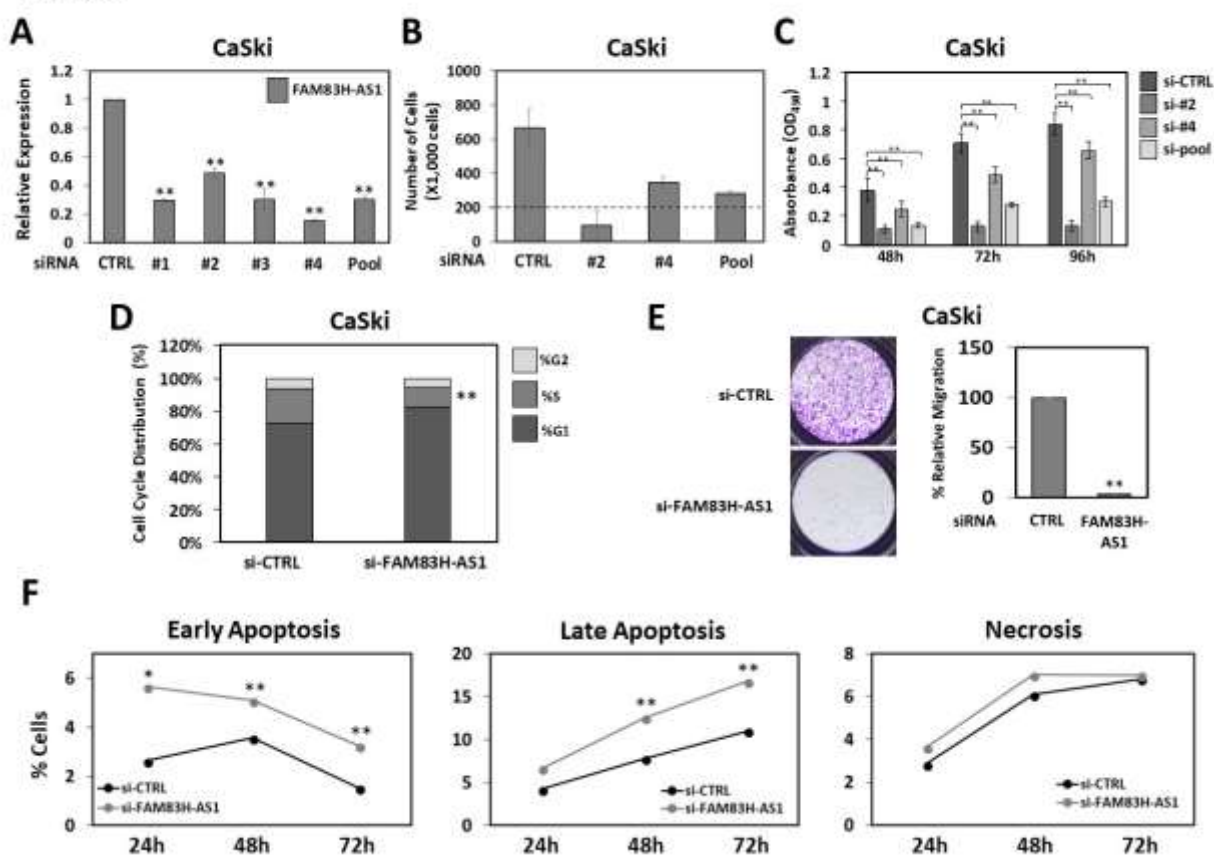
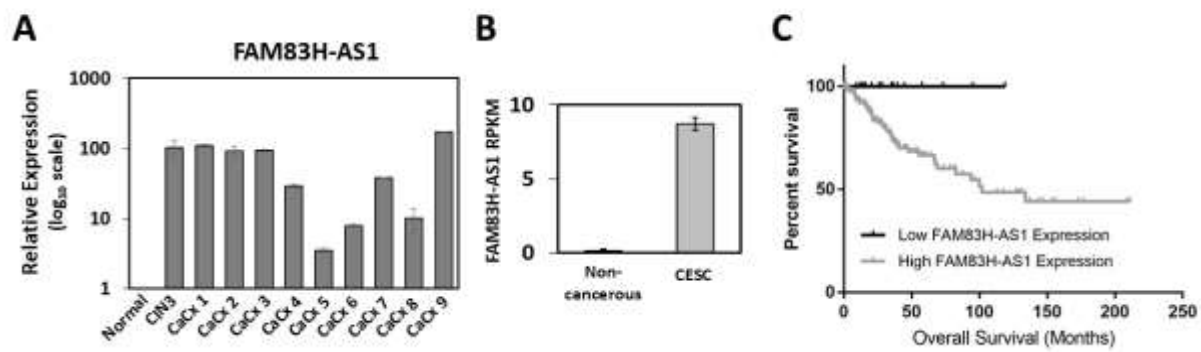


Figure 6:



XI. Supplementary Figures

Supplementary Figure S1

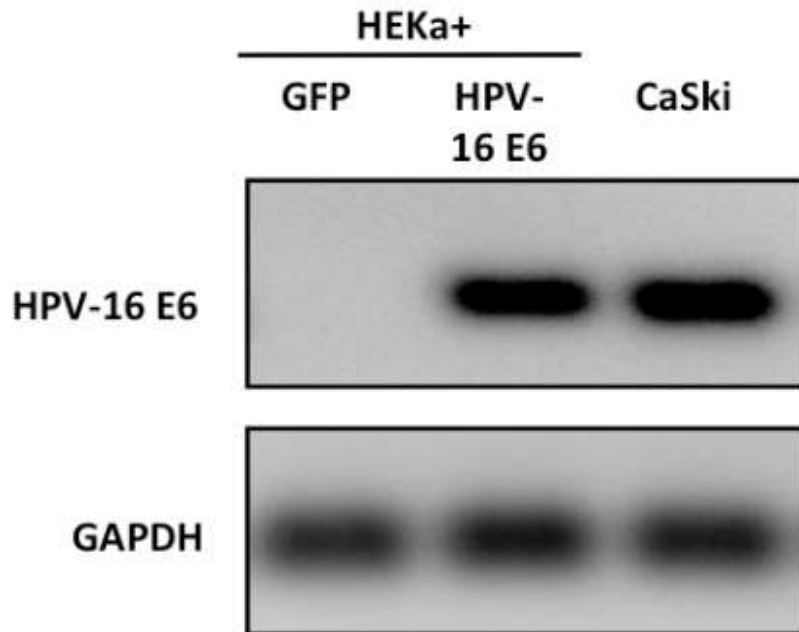


Figure S1: Validation of stable HPV-16 E6 expression in primary foreskin keratinocytes. RT-PCR was used to confirm stable HPV-16 E6 expression in primary foreskin keratinocytes (HEKa) compared to HEKa expressing GFP control and HPV-16 positive CaSki cell line.

Supplementary Figure S2

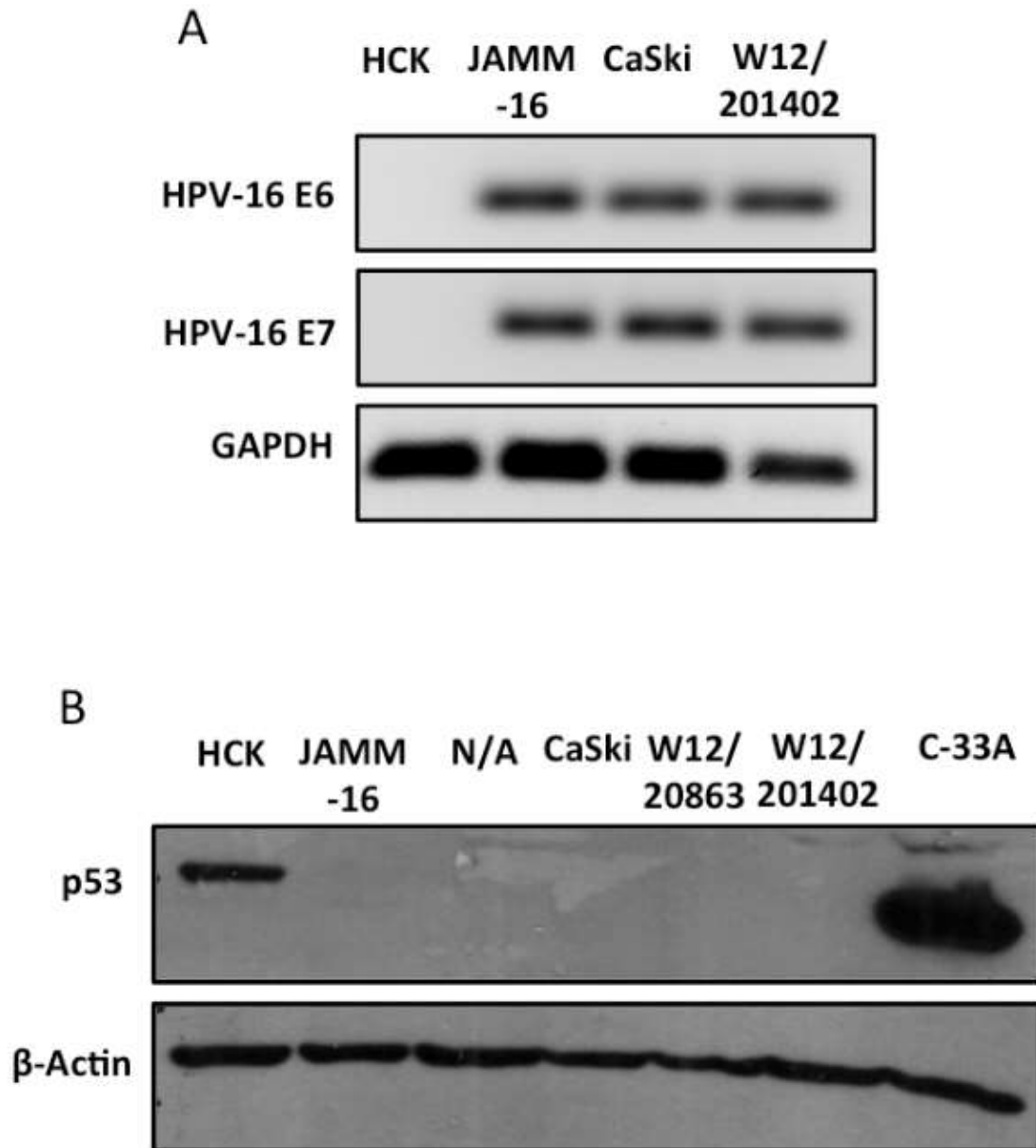


Figure S2: Validation of HPV-16 expression in primary cervical keratinocytes. RT-PCR was used to confirm HPV-16 E6 and E7 expression in cervical keratinocytes transfected with entire HPV-16 genome (JAMM-16) compared to uninfected primary cervical keratinocytes (HCK) and HPV-16 positive CaSki and W12/201402 cell lines. GAPDH was used as loading control. (B) Western blot analysis was used to confirm p53 degradation in HPV-16 positive JAMM-16, CaSki, W12/20863, W12/201402 cell lines compared to HCK and HPV negative, p53-mutated C-33A cell line. Note: N/A represents a cervical cell line that was not used in this study.

Supplementary Figure S3

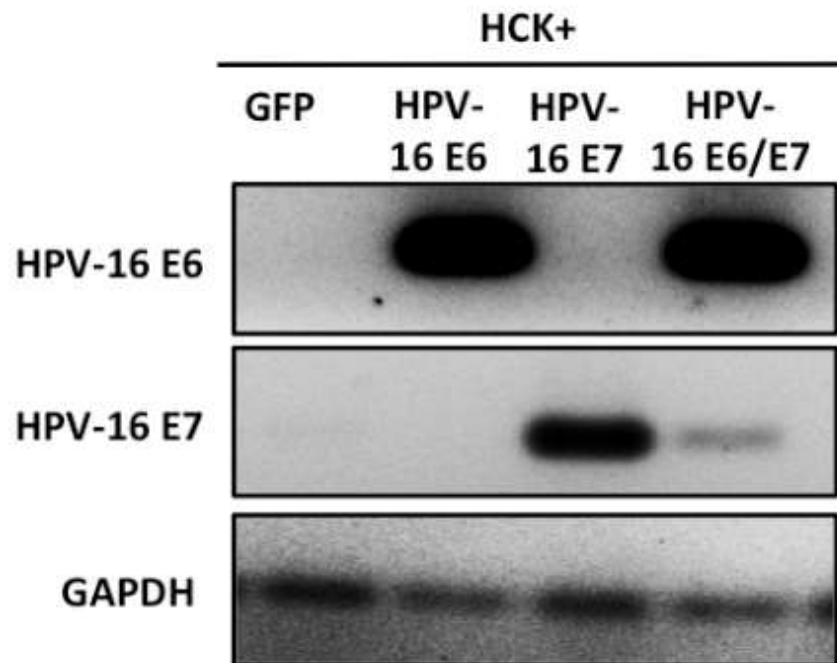
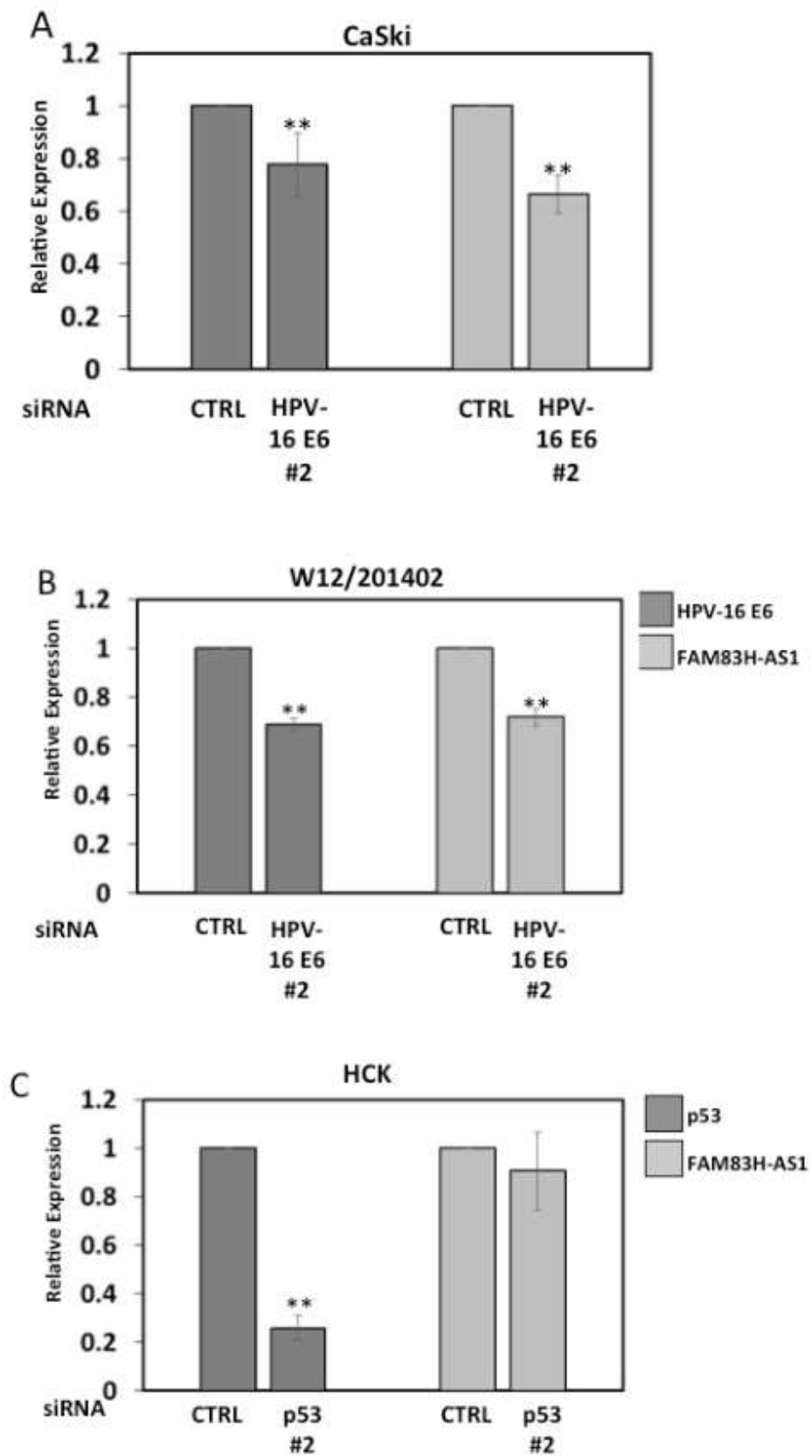


Figure S3: Validation of stable HPV-16 E6 and E7 expression in primary cervical keratinocytes. RT-PCR was used to confirm stable HPV-16 E6 and E7 expression, as desired, in primary cervical keratinocytes (HCK) compared to HCK expressing GFP control.

Supplementary Figure S4



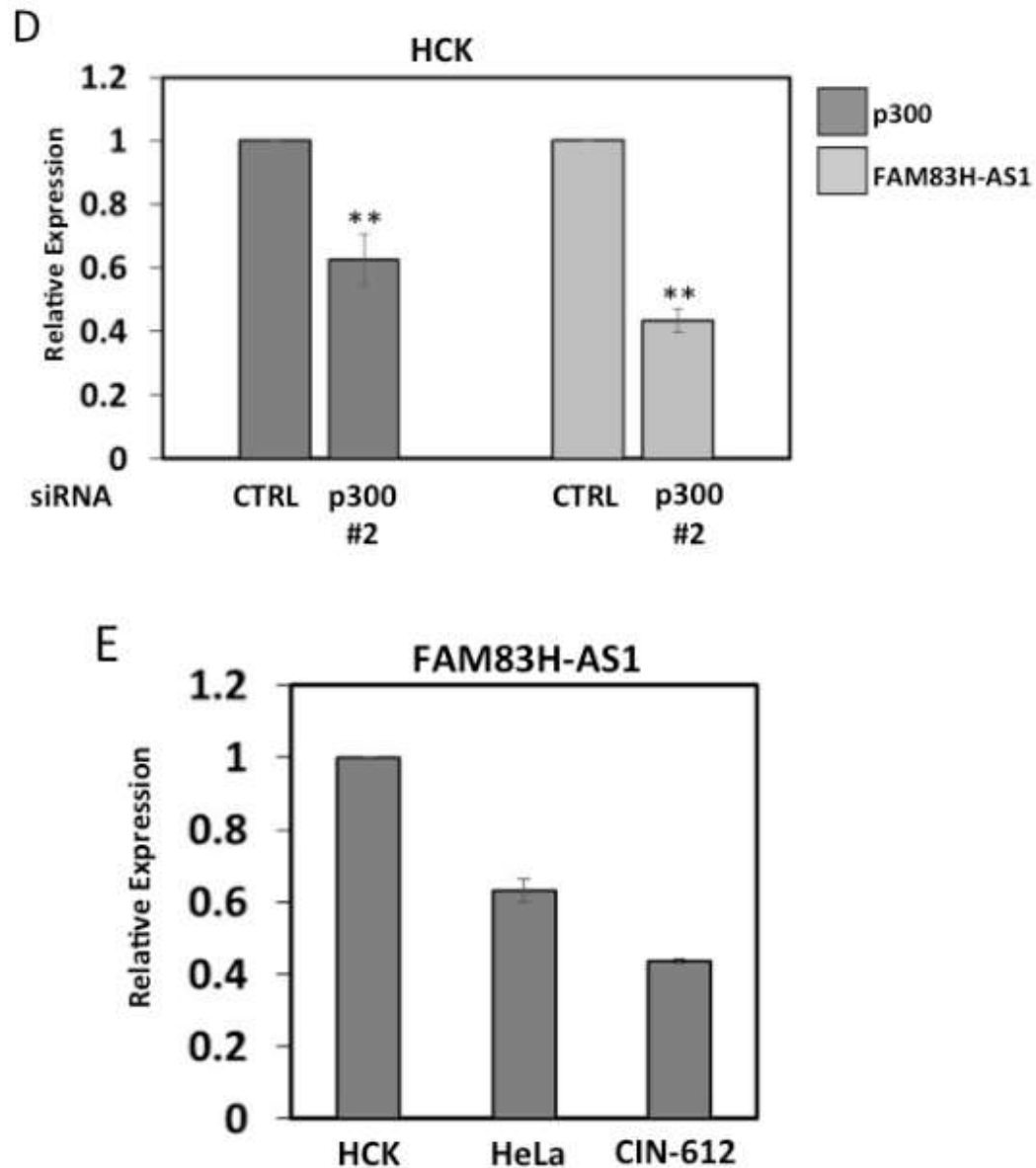


Figure S4: FAM83H-AS1 expression is regulated by HPV-16 E6 in a p53-independent, p300-dependent manner. (A-B) qRT-PCR analysis of HPV-16 E6 and FAM83H-AS1 expression in HPV-16 positive Caski (A) and W12/201402 (B) cell lines transfected with a second siRNA (siRNA #2) against HPV-16 E6 compared to siRNA control. (C) qRT-PCR analysis of p53 and FAM83H-AS1 expression in primary cervical keratinocytes (HCK) transfected with a different siRNA (siRNA #2) against p53 compared to siRNA control. (D) p300 and FAM83H-AS1 expression in HCK transfected with a different siRNA (siRNA #2) against p300 compared to siRNA control. (E) qRT-PCR analysis of FAM83H-AS1 expression in HPV-18 positive HeLa cells and HPV-31b positive CIN-612 cervical cell lines compared to primary cervical keratinocytes (HCK). (A-D) The graphs show the average of two independent experiments. Similar results were obtained in at least three independent experiments. GAPDH mRNA was used to normalize the qRT-PCR analyses. Two-tailed t test results are indicated as ** $p \leq 0.01$. CTRL, control.

Supplementary Figure S5

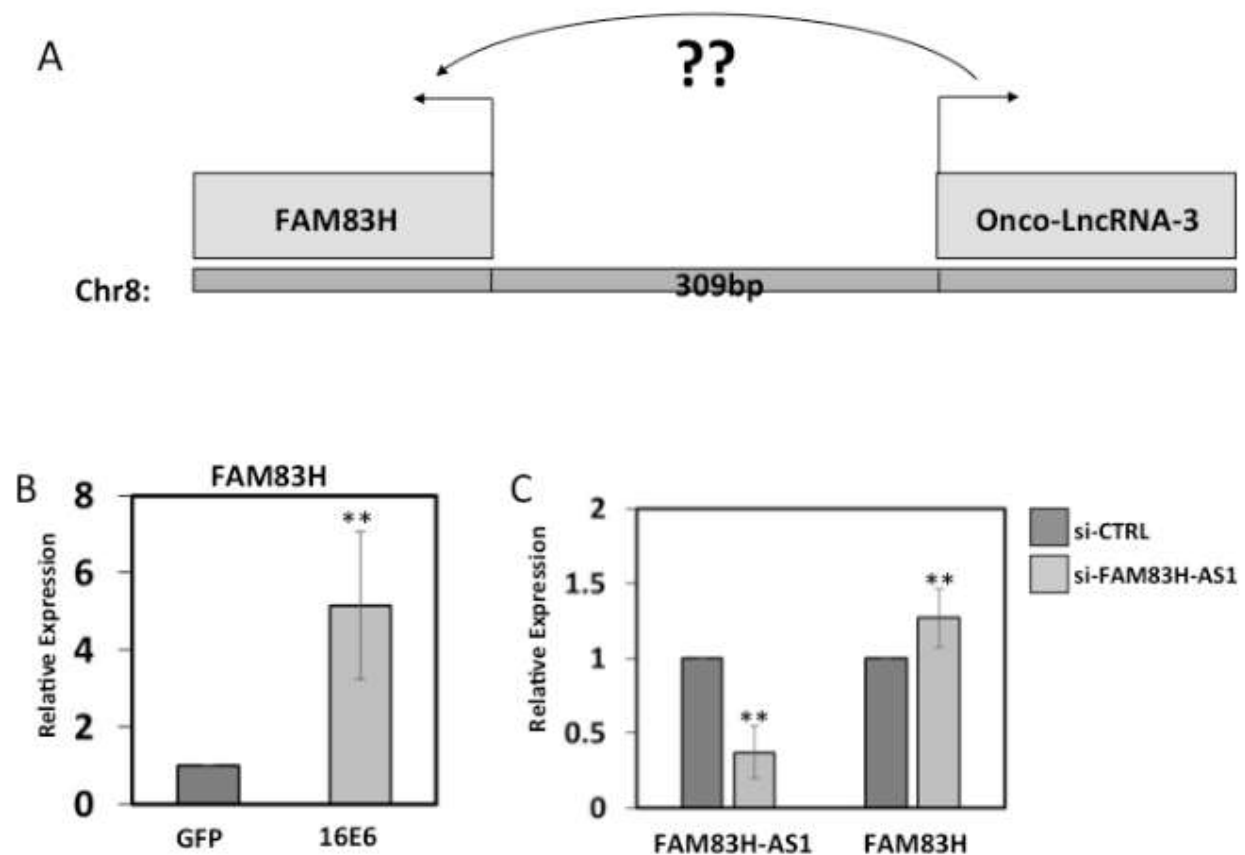


Figure S5: FAM83H expression is not regulated by FAM83H-AS1. (A) Schematic diagram depicting possible cis-regulation of protein coding gene FAM83H by lncRNA FAM83H-AS1. (B) qRT-PCR analysis of FAM83H protein coding gene expression in primary cervical keratinocytes (HCK) stably expressing HPV-16 E6 compared to GFP control. (C) qRT-PCR analysis of FAM83H and FAM83H-AS1 expression in CaSki transfected with siRNA against FAM83H-AS1. The graphs shows the average of two independent experiments. Similar results were obtained in at least three independent experiments. GAPDH mRNA was used to normalize the qRT-PCR analyses. Two-tailed t test results are indicated as ** $p \leq 0.01$. CTRL, control.

Supplementary Figure S6

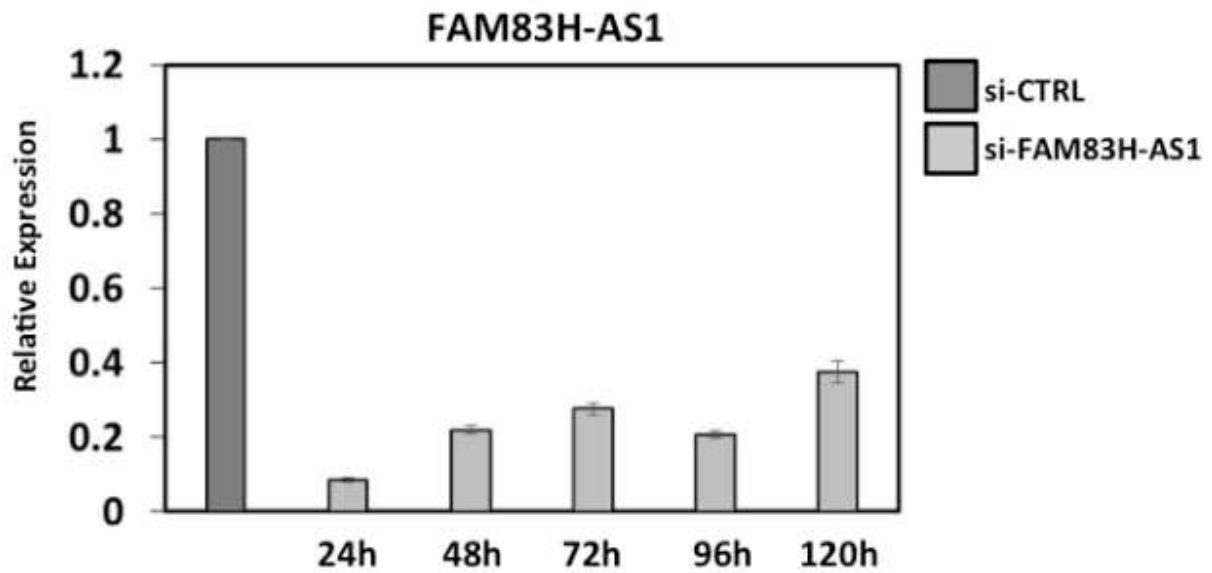
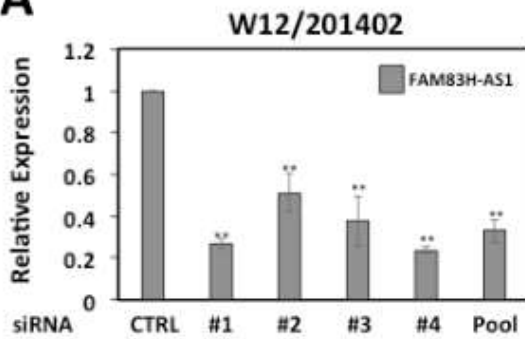


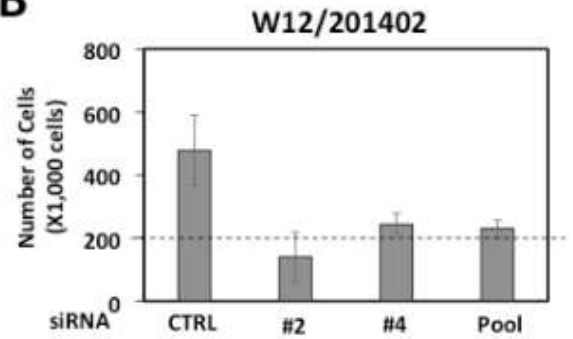
Figure S6: FAM83H-AS1 expression at 24, 48, 72, 96, and 120 hours post-transfection with FAM83H-AS1 siRNA SMARTpool by qRT-PCR analysis.

Supplementary Figure S7

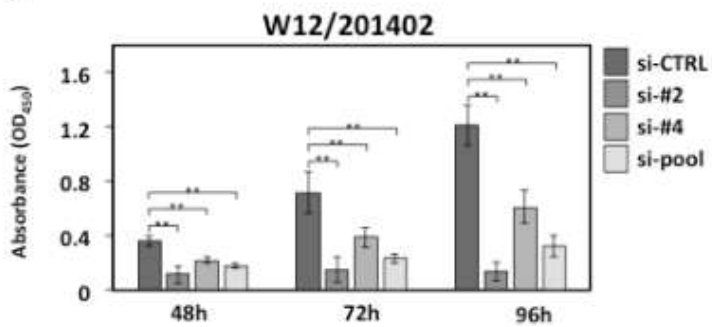
A



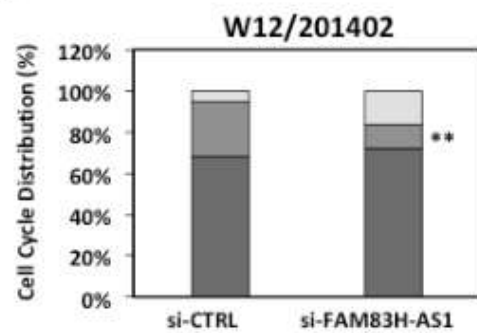
B



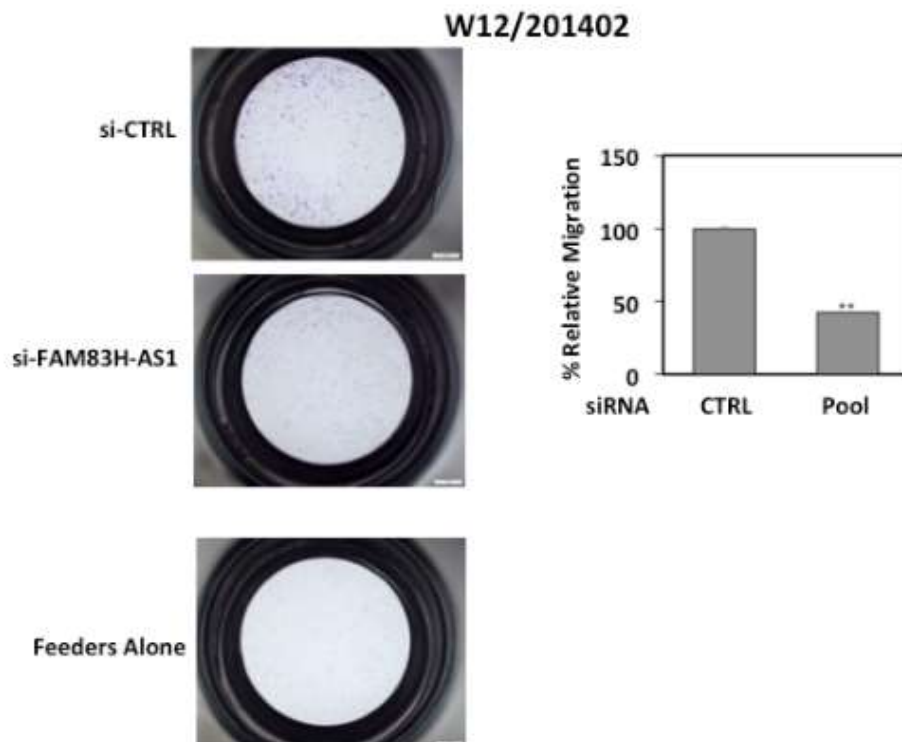
C



D

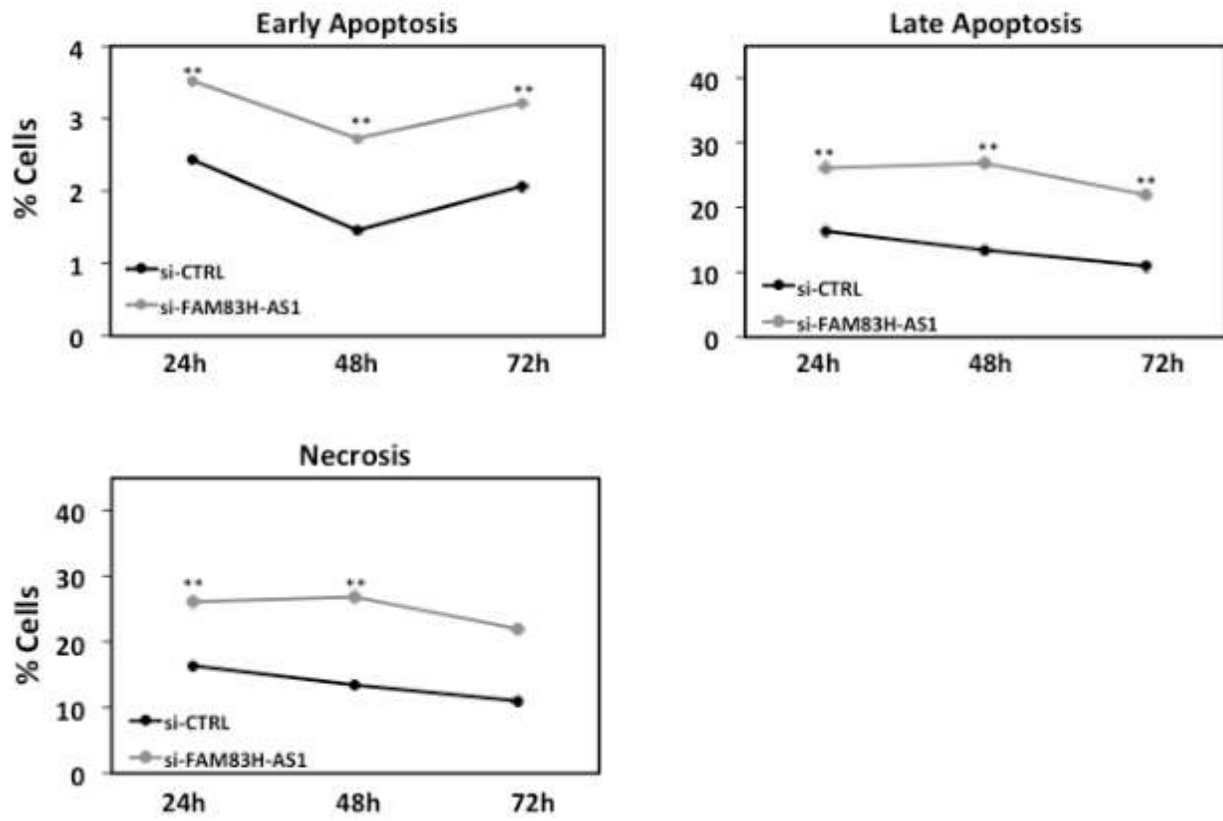


E



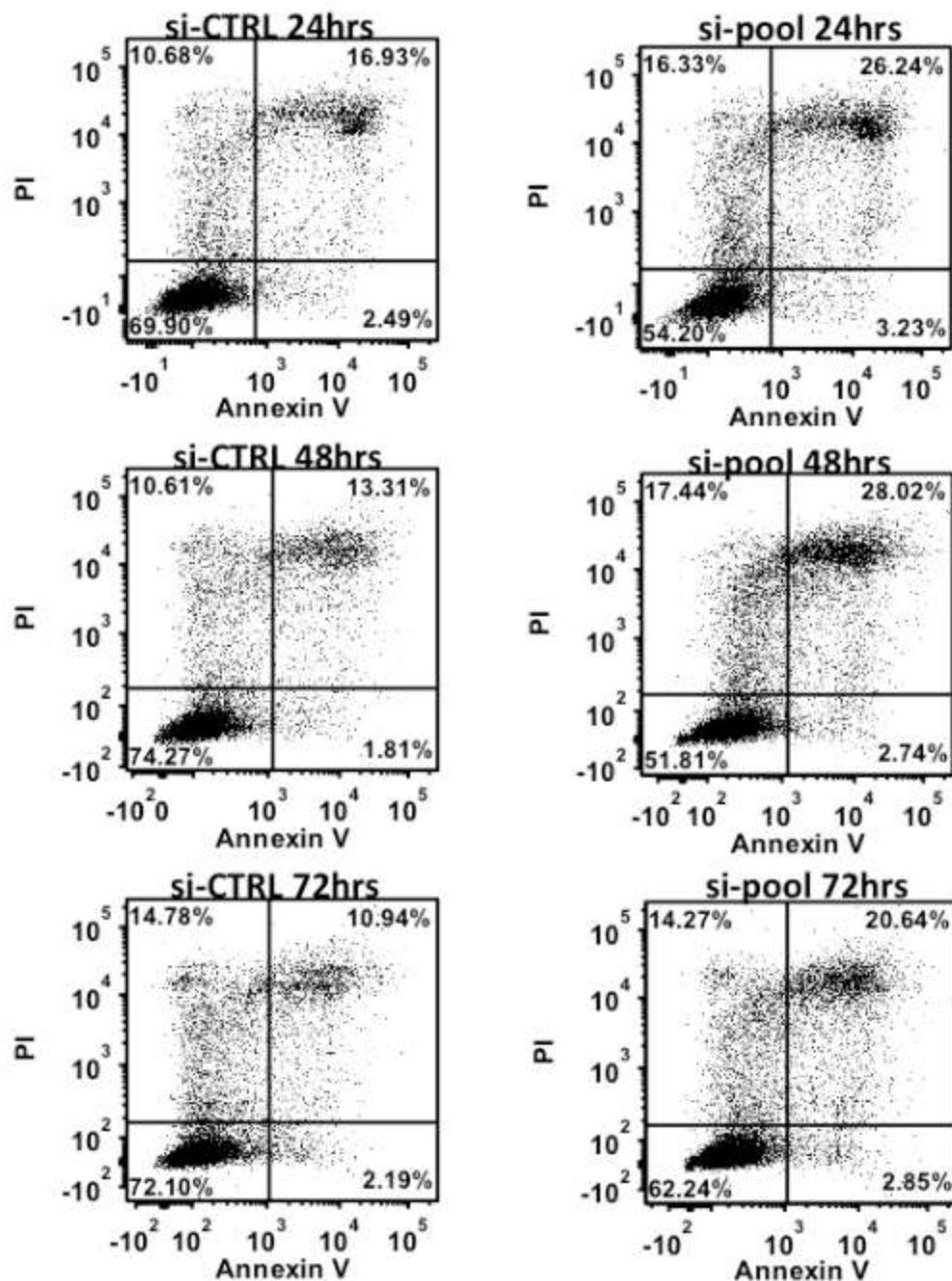
F

W12/201402



G

201402



H

CaSki

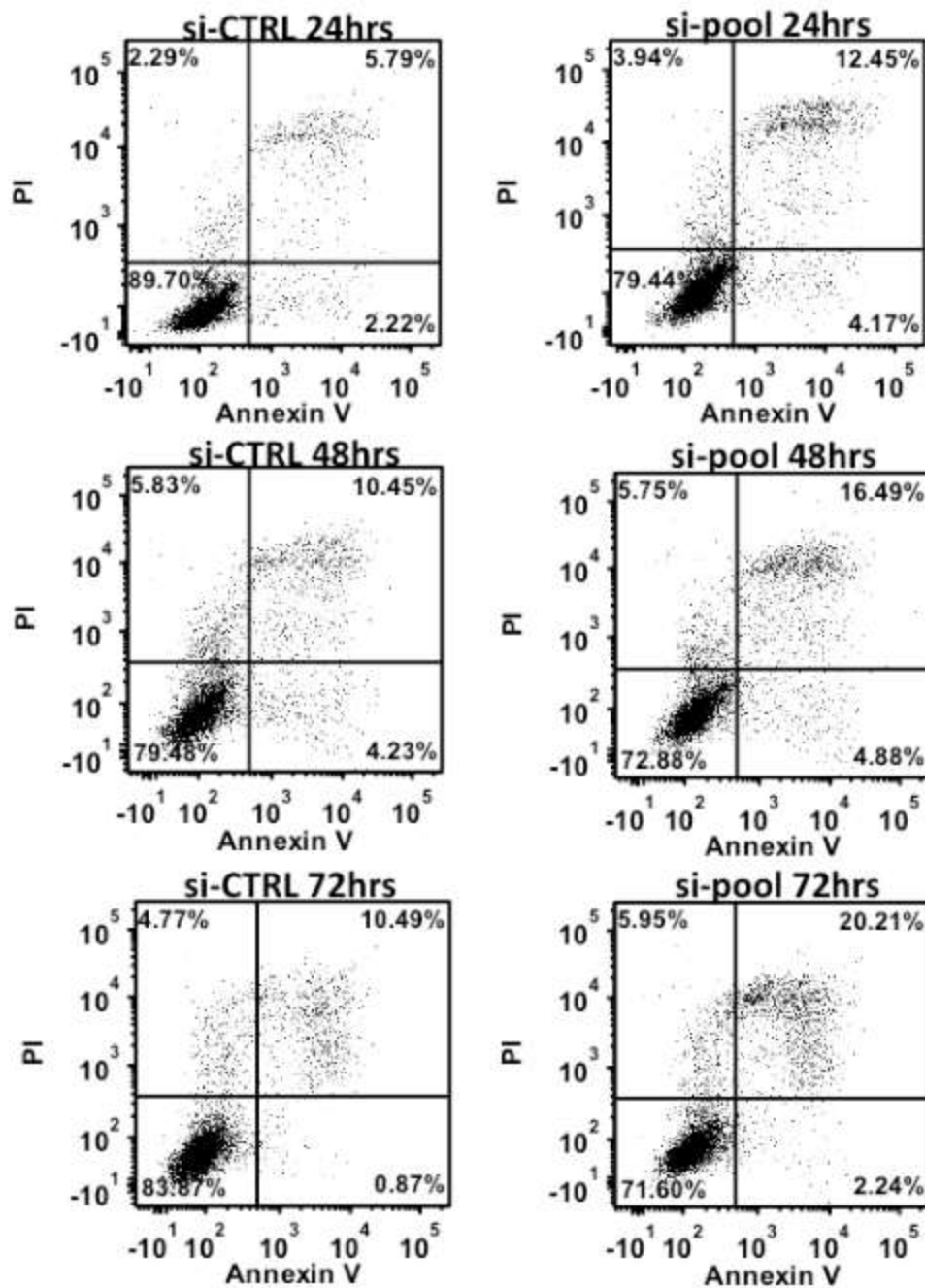


Figure S7: FAM83H-AS1 knockdown altered cell proliferation, migration, and apoptosis in W12/201402 cells. (A) Knockdown efficiency of individual and SMARTpool siRNA against FAM83H-AS1 in HPV-16 positive W12/201402, measured by qRT-PCR analysis. Because of variations in the expression of GAPDH after the knockdown of FAM83H-AS1, we used GUSB mRNA to normalize the qRT-PCR analyses. The graph shows average of two individual experiments. (B) W12/201402 cells were transfected with individual siRNAs against FAM83H-AS1, siRNA SMARTpool against FAM83H-AS1, or siRNA control for 48 hours. Cells were then re-plated in equal numbers (200,000 cells/well, represented by dashed line in graph) and cultured another 48 hours prior to re-counting attached cells. The graph shows the average of two individual experiments. (C) W12/201402 cells were transfected with individual siRNAs against FAM83H-AS1, siRNA SMARTpool against FAM83H-AS1, or siRNA control for 48 hours then plated in equal numbers. Transfected cells were analyzed for cellular proliferation assessment by CCK-8 assay at 48, 72, and 96 hours post-plating. The graph shows the average of two individual experiments. (D-G) W12/201402 cells were transfected with siRNA SMARTpool against FAM83H-AS1 or siRNA control for 24 hours then plated in equal numbers for experiments. (D) Transfected cells were analyzed for cell cycle alterations by FACS analysis. W12/201402 cells with knockdown of FAM83H-AS1 exhibit less cells in S-phase of cell cycle compared to control cells. The graph shows the average of two individual experiments. (E) Transwell migration of transfected cells was analyzed 48 hours post-plating in upper chamber with chemoattractant in lower chamber. The graph shows the average of two individual experiments. (F-H) Transfected W12/201402 or CaSki cells were collected at 1, 2, and 3 days post-plating, stained with Annexin V/PI, and analyzed by flow cytometry to show alterations in apoptosis compared to siRNA control. (F) Quantitative analysis of Annexin V/PI flow cytometry data in W12/201402 cells. The graph shows the average of two individual experiments. (G) Dot plot analysis of Annexin V/PI flow cytometry data in W12/201402 cells. (H) Dot plot analysis of Annexin V/PI flow cytometry data in CaSki cells. Representative images shown for siRNA FAM83H-AS1 and siRNA control at each time point. Two-tailed t test results are indicated as * $p \leq 0.05$ and ** $p \leq 0.01$.

XII. Supplementary Tables

Table S1: Host lncRNAs altered after expression of HPV-16 E6 in primary foreskin keratinocytes.

Table shows details for each host lncRNA that was altered 1.5-fold or greater in primary foreskin keratinocytes (HEKa) expressing HPV-16 E6 compared to uninfected HEKa. Triplicates for each sample were sent for high-throughput RNA sequencing analysis.

Gene ID	Fold Change ([HPV 16E6] vs [Control])	Gene Symbol	Chromosome	Start	End	Strand	RPKM GFP Control -1	RPKM GFP Control -2	RPKM GFP Control -3	RPKM HPV-16 E6-1	RPKM HPV-16 E6-2	RPKM HPV-16 E6-3
ENSG00000130600	4.99	H19	chr11	2016406	2022700	-	2.1290493	2.1426256	5.6363993	19.229809	10.783182	15.414708
ENSG00000254236	3.54	KB-1639H6.2	chr8	104032415	104033656	-	0	0	0.2873624	3.250996	3.96886	3.4498966
ENSG00000244528	3.31	AC134873.1	chr2	243064438	243064620	+	1.2767934	2.8750565	2.7582934	6.186408	6.602485	9.019768
ENSG00000273413	3.07	RP11-96C23.15	chr10	88729994	88730583	-	8.393546	0	0	4.7901816	7.0136943	7.2466874
ENSG00000228626	3.06	RP11-495P10.9	chr1	147760107	147761057	-	2.0288084	0.88392633	2.2444494	6.9817734	5.0385094	3.6965914
ENSG00000230937	2.94	MIR205HG	chr1	209602165	209606183	+	61.548264	49.323284	62.656292	218.4987	181.15729	121.917755
ENSG00000254860	2.73	TMEM9B-AS1	chr11	8986222	8999074	+	0.8062374	1.7021289	0.6363604	4.052271	3.1377754	2.730444
ENSG00000249641	2.69	HOXC-AS5	chr12	54329112	54333427	-	0.9593882	1.8041867	0.9305784	3.716362	3.7588704	2.5167224
ENSG00000270168	2.56	LA16c-380H5.4	chr16	3051301	3052017	+	6.4768987	0.8551222	4.0893016	9.19451	7.0534296	6.8577585
ENSG00000235314	2.53	LINC00957	chr7	44078770	44081905	+	0.35596678	0.73328674	0.70044947	1.9380348	2.741936	3.0452104
ENSG00000230409	2.50	TCEA1P2	chr3	37317087	37318089	+	7.469508	9.592129	10.218497	22.008438	24.913902	20.846043
ENSG00000233223	2.45	AC113189.5	chr17	7485282	7487390	-	1.4596063	2.9885232	2.0333843	3.720071	5.5129123	6.3233705
ENSG00000254192	2.36	CTC-558O2.2	chr5	168081518	168094766	+	1.2354965	0.90355885	1.6458715	3.684764	2.9198492	2.4772859
ENSG00000251580	2.35	RP11-539L10.3	chr4	6672452	6675557	-	0.83309	1.3336157	1.0557044	2.159164	2.0041137	4.219327
ENSG00000220804	2.29	AC093642.5	chr2	243030784	243102304	+	1.022857	1.6583912	2.8178344	4.2308583	2.5898807	5.2071075
ENSG00000201428	2.28	RN7SKP71	chr12	112704881	112705198	+	0.05682099	0.26406604	0.33525583	0.033641126	0.46484584	11.79032
ENSG00000238107	2.27	RP11-495P10.5	chr1	147767287	147769663	+	1.1964403	0.5992599	0.778185	3.4525478	1.8964818	2.1271327
ENSG00000228463	2.25	AP006222.2	chr1	227615	267253	-	0.92179984	1.8147042	2.4698875	3.840908	3.5203686	3.7663238
ENSG00000263934	2.22	SNORD3A	chr17	19091329	19092027	+	56.559566	96.54193	87.400215	115.812065	101.63637	442.51657
ENSG00000267984	2.21	CTD-2616J11.9	chr19	51848423	51856547	+	0.039008502	0	0	0.09365486	0	10.77701
ENSG00000200488	2.18	RN7SKP203	chr2	76672205	76672536	-	0.3945807	2.4766135	0.891996	1.1814926	0.6042596	21.76185
ENSG00000255717	2.13	SNHG1	chr11	62619460	62623386	-	7.0900187	3.8827178	5.670217	13.077036	11.900302	9.669952
ENSG00000259001	2.12	RPPH1	chr14	20811207	20811844	-	15.001923	52.59274	31.222376	27.470306	36.77239	231.42776
ENSG00000234741	2.10	GAS5	chr1	173833038	173838020	-	75.5498	34.682	63.601	141.4	116.2	93.72

							6	1	818	7461	6347	997
ENSG00000215039	2.08	CD27-AS1	chr12	6548167	6560733	-	0.66829 19	0.8538 55	0.8496 1987	3.074 375	1.768 162	1.665 7759
ENSG00000259970	2.06	AC099668.5	chr3	49721913	49722416	-	0.57804 507	1.3075 932	0.8779 387	2.843 681	2.354 5742	1.696 4713
ENSG00000254578	2.04	CTD- 2517M22.16	chr8	145689200	145690484	+	2.56431 48	1.6997 473	2.1004 875	4.821 358	4.543 132	3.554 0457
ENSG00000202198	2.04	RN7SK	chr6	52860418	52860748	+	77.7898 2	202.92 421	175.02 835	174.8 5051	183.0 6963	733.0 992
ENSG00000255831	2.04	AL139385.1	chr13	111291555	111292340	+	0	2.5623 991	0	0	5.166 7576	4.200 331
ENSG00000271992	2.02	RP11-42O15.3	chr1	70910754	70911219	+	0.34897 357	0.0900 998	0.8007 2904	2.295 6822	1.722 0101	2.083 4737
ENSG00000227195	2.01	MIR663A	chr20	26167556	26232162	-	0.89503 71	1.0255 735	1.4251 992	1.545 9784	1.247 556	6.125 4115
ENSG00000239002	1.98	SCARNA10	chr12	6619388	6619717	+	0.54754 776	1.9084 775	1.3999 47	1.167 0413	2.879 6296	6.134 6283
ENSG00000263917	1.97	RP11-53I6.2	chr18	29598792	29691742	+	1.69259 02	0.7714 8676	1.5012 345	1.865 241	3.775 4347	2.767 804
ENSG00000266929	1.96	RP11- 400F19.8	chr17	40688528	40714080	+	2.64025 28	2.2248 166	2.0022 075	4.118 1316	4.623 1112	4.685 91
ENSG00000223891	1.96	OSER1-AS1	chr20	42839600	42854667	+	0.42706 716	1.3540 938	1.7394 654	2.074 9772	3.062 2957	2.806 612
ENSG00000244879	1.96	GABPB1-AS1	chr15	50647525	50659636	+	1.49664 08	1.0193 752	1.1708 221	2.835 0017	2.036 269	2.337 8665
ENSG00000225976	1.96	RP11- 192N10.2	chr10	29986832	29987157	+	4.10156 9	1.4167 225	3.0522 678	4.856 7057	5.441 2627	5.069 306
ENSG00000259827	1.95	RP11- 343H19.2	chr16	56716382	56721719	+	3.48532 18	2.4112 535	2.5835 383	3.920 711	3.862 1109	10.60 9867
ENSG00000235499	1.95	AC073046.25	chr2	74212259	74213470	+	4.49410 3	3.1759 024	2.5058 224	6.528 4495	5.577 129	7.256 5813
ENSG00000258377	1.94	RP11-649E7.5	chr14	50087533	50090198	-	0	3.7360 387	4.1599 88	5.070 155	4.762 3515	4.735 923
ENSG00000233825	1.94	RP11- 135A24.4	chr10	32635427	32636107	+	2.00523 7	0.9287 9003	1.8029 468	2.122 1492	3.477 26	3.551 1832
ENSG00000269926	1.93	RP11- 442H21.2	chr10	74034673	74035738	-	37.1678 16	17.135 89	27.302 435	56.37 6045	53.58 8455	41.68 311
ENSG00000261824	1.91	LINC00662	chr19	28175488	28284848	-	1.77218 08	0.9460 5726	1.4761 61	3.744 9877	2.442 1446	1.996 9538
ENSG00000267317	1.90	CTB-25B13.12	chr19	1457664	1462764	-	0.77052 76	2.7577 708	2.0533 23	3.240 5212	3.709 7907	3.223 024
ENSG00000179859	1.89	AC025335.1	chr17	7816642	7819271	-	2.48756 43	3.6358 418	4.2272 625	7.583 128	6.336 298	5.399 1914
ENSG00000257653	1.87	RP11-579D7.2	chr12	49159977	49161106	+	0.70771 77	1.0421 093	1.3791 414	2.705 9877	1.943 5432	1.797 2413
ENSG00000238271	1.87	IFNWP19	chr9	21455483	21456048	+	6.09118 27	4.1468 25	6.1276 43	10.78 6538	8.897 233	10.57 8037
ENSG00000224032	1.87	EPB41L4A-AS1	chr5	111496223	111499973	+	7.39583 83	3.9629 142	7.1221 61	9.977 508	12.22 6742	11.20 5975
ENSG00000257270	1.87	RP11- 521B24.5	chr14	105934130	105936954	-	6.40190 6	3.2108 22	3.5815 47	8.843 35	7.365 5524	7.390 7046
ENSG00000257084	1.86	U47924.27	chr12	7072409	7073610	+	31.198	36.405 44	20.322 99	53.07 6824	61.99 1566	44.91 036
ENSG00000225377	1.85	RP5-1103G7.4	chr20	300957	328868	-	2.56064 6	3.4069 042	5.1782 656	6.993 363	5.695 6306	7.200 583
ENSG00000239470	1.84	RP11-16F15.2	chr11	9681946	9682503	+	6.97289 23	3.0191 91	4.1343 21	10.61 4801	8.628 567	5.930 6784
ENSG00000198496	1.84	NBR2	chr17	41277627	41305688	+	0.83190 143	1.0633 509	1.4072 425	2.247 7791	2.288 2934	1.808 5096
ENSG00000203499	1.84	FAM83H-AS1	chr8	144816310	144828507	+	9.59697 8	3.3299 105	4.7020 435	12.95 5384	9.471 8895	7.570 58
ENSG00000177410	1.83	ZFAS1	chr20	47894715	47905797	+	211.102 16	139.73 38	252.70 905	418.7 3697	367.0 7626	297.5 4745

ENSG00000224891	1.83	AC007899.3	chr2	37375673	37376447	+	4.01025 72	2.9299 183	3.7727 127	6.872 386	6.110 306	6.431 074
ENSG00000214293	1.82	RSBN1L-AS1	chr7	77286977	77325582	-	1.88850 9	1.6980 832	1.8871 909	3.833 9822	3.262 9604	2.897 0351
ENSG00000250902	1.81	RP11- 301H24.4	chr4	146435767	146438270	-	6.98976 66	3.0939 543	5.3921 43	10.33 7118	8.141 412	8.172 936
ENSG00000259375	1.81	RP11- 815J21.2	chr15	86122277	86123409	-	1.83370 24	1.1494 807	1.1902 504	1.597 2154	2.884 9528	3.210 236
ENSG00000012171	1.80	SEMA3B	chr3	50304990	50314977	+	0.94935 43	0.3429 1044	1.0639 977	2.247 6633	1.790 5937	1.540 2615
ENSG00000273344	1.79	PAXIP1-AS1	chr7	154795158	154797413	+	1.07325 18	2.6334 62	2.4573 538	3.006 4073	3.060 8828	4.358 569
ENSG00000260552	1.79	RP11-49I11.1	chr18	33759959	33767411	-	1.30548 33	2.1372 445	4.7266 11	3.969 9843	6.068 1834	3.154 5815
ENSG00000243004	1.79	AC005062.2	chr7	19958604	20180076	-	4.87558 13	7.2985 44	3.9218 314	6.931 1867	14.13 0893	8.209 532
ENSG00000196756	1.78	SNHG17	chr20	37049235	37063996	-	2.54244 57	1.8823 376	4.6329 513	4.370 9865	5.745 9593	5.005 6214
ENSG00000272906	1.78	RP11- 533E19.7	chr1	179850742	179851730	-	0.29894 31	1.2788 76	1.0368 878	1.577 8978	1.703 776	2.794 0536
ENSG00000261771	1.78	DYX1C1- CCPG1	chr15	55647446	55790558	-	2.09689 38	2.6955 016	2.8982 72	5.290 535	4.608 052	3.804 3263
ENSG00000204054	1.78	LINC00963	chr9	132245730	132275947	+	1.01223 13	0.8181 0194	0.7509 8056	2.690 7895	1.890 5445	1.125 3008
ENSG00000269893	1.78	SNHG8	chr4	119199864	119200978	+	95.6830 75	85.640 51	110.78 4874	193.3 1099	208.9 9217	126.6 07704
ENSG00000269900	1.78	RMRP	chr9	35657748	35658015	-	61.7190 86	117.54 486	105.27 629	99.84 465	121.5 6224	352.0 1825
ENSG00000265735	1.76	RN7SL5P	chr9	9442060	9442347	+	89.3426 7	171.01 642	175.36 769	146.4 969	160.5 5365	623.5 067
ENSG00000229413	1.75	AC018638.1	chr7	128293740	128293989	-	2.08112 76	0.8533 8426	0.6139 542	2.203 358	1.459 3241	3.445 703
ENSG00000250299	1.74	MRPS31P4	chr13	53191693	53211581	+	0.99767 58	0.6966 341	0.8349 367	2.199 7466	1.885 6709	1.279 5835
ENSG00000226419	1.74	RP11-31F15.1	chr1	113499037	113542118	+	3.05203 82	2.3166 275	2.4479 396	5.745 393	4.417 4857	3.588 944
ENSG00000245060	1.74	LINC00847	chr5	180257957	180262726	+	2.51798 3	1.9512 683	1.9726 743	3.667 7196	3.663 7437	3.780 457
ENSG00000110811	1.73	LEPREL2	chr12	6937572	6949018	+	1.43181 74	1.6433 642	1.4975 213	3.065 626	2.159 1876	2.762 544
ENSG00000262160	1.73	RP11-96D1.11	chr16	68259872	68263048	-	1.01642 31	1.7172 104	1.3368 715	2.896 4238	2.114 6343	1.976 8672
ENSG00000250251	1.73	PKD1P6	chr16	15219099	15248421	-	4.10675 76	2.8266 087	4.4332 04	4.759 571	5.149 509	10.82 3436
ENSG00000203512	1.72	AL353662.2	chr9	34195642	34195883	+	1.73161 97	0.8091 1475	1.1475 528	3.164 7897	1.525 1368	2.108 7782
ENSG00000227615	1.72	RP11-864N7.2	chr11	74456761	74457159	-	6.18152 57	1.2890 593	4.9579 773	6.729 743	6.448 0925	4.633 6823
ENSG00000257086	1.72	RP11- 783K16.13	chr11	64014411	64016966	-	1.58257 52	1.1539 651	1.8851 668	3.347 7898	2.183 432	2.384 6471
ENSG00000230795	1.72	HLA-K	chr6	29894236	29897009	+	0.61804 247	1.1913 046	1.5108 755	1.330 7873	2.285 3572	2.992 5754
ENSG00000180672	1.71	AC007362.1	chr2	206642540	206644433	+	1.37003 06	0.6789 9704	1.1776 617	3.312 5985	1.914 6636	1.281 505
ENSG00000254682	1.71	RP11- 660L16.2	chr11	71159720	71163203	+	1.14260 32	0.7100 659	0.4964 7444	1.667 6105	1.925 3993	1.792 3354
ENSG00000267023	1.71	LRRC37A16P	chr17	66121918	66148609	-	1.45405 14	1.6921 17	1.4601 423	2.523 746	2.944 7954	2.423 118
ENSG00000227081	1.71	RP11- 543P15.1	chr12	3320775	3321096	-	7.24820 5	1.5755 806	5.4170 194	7.375 556	8.514 681	4.918 4346
ENSG00000223745	1.71	RP4-717I23.3	chr1	93727743	93811582	-	0.95840 377	1.4749 708	1.4199 601	2.833 7555	1.944 4739	1.893 2604
ENSG00000226752	1.70	PSMD5-AS1	chr9	123587106	123616651	+	0.90399	0.6311	1.0164	2.319	1.464	1.481

							02	066	516	7796	0397	2256
ENSG00000227632	1.70	AC018804.6	chr2	130959885	130970201	+	1.50575 63	2.2844 598	2.4223 318	2.553 2963	4.621 544	3.451 2596
ENSG00000228327	1.69	RP11- 206L10.2	chr1	700237	714006	-	3.12584 38	1.0653 372	3.5199 823	3.532 1116	3.393 6765	4.746 0814
ENSG00000228638	1.69	FCF1P2	chr3	48332283	48332865	-	0.71284 515	0.7201 8015	0.2133 4462	2.789 1552	1.738 6442	0.814 9583
ENSG00000258186	1.69	SLC7A5P2	chr16	21531151	21531686	-	8.26003	6.4758 79	14.699 861	13.65 3822	13.00 4153	21.40 1606
ENSG00000225648	1.68	SBDSP1	chr7	72300004	72307909	+	10.7959 62	11.587 327	13.133 697	19.68 8213	19.81 655	19.96 0356
ENSG00000272711	1.68	RP11- 259N19.1	chr2	75059782	75061114	-	1.45040 58	1.1969 146	1.6662 19	3.884 3012	2.328 7628	1.503 2032
ENSG00000256462	1.67	RP11-116G8.5	chr10	5566939	5567705	-	3.94108 49	0	0	3.097 3866	5.978 923	0
ENSG00000236439	1.67	RP11-175B9.3	chr1	202440937	202441258	-	6.70856 76	3.0561 595	4.4550 18	7.306 8976	8.068 728	7.226 4347
ENSG00000233016	1.67	SNHG7	chr9	139615818	139622636	-	3.80191 66	2.5988 54	3.9963 634	4.894 757	6.888 8636	5.446 209
ENSG00000234975	1.67	FTH1P2	chr1	228823162	228823574	+	2.36462 16	8.7332 3	5.5768 6	7.877 958	8.395 835	8.092 968
ENSG00000228998	1.66	RP11-697E2.7	chr15	90818266	90820841	+	1.66241 1	1.0105 448	1.4692 17	2.458 5187	2.098 612	2.181 2184
ENSG00000251022	1.66	THAP9-AS1	chr4	83814162	83822113	-	2.55584 17	2.0884 433	4.3881 38	5.688 368	4.508 133	4.144 217
ENSG00000256633	1.64	RP11-169D4.2	chr11	72295616	72299023	+	0.36736 822	0.2219 0487	0.2809 7868	0.585 01613	1.582 57	2.773 093
ENSG00000266754	1.64	RN7SL524P	chr10	105734277	105734572	-	0.67657 28	0.7092 3144	1.0905 253	2.662 4227	0.618 2988	1.791 2405
ENSG00000260269	1.63	CTD- 2323K18.1	chr15	75819491	75893546	-	1.65046 27	0.9741 242	0.6995 654	3.534 106	1.676 849	1.202 5416
ENSG00000260088	1.62	RP11-92G12.3	chr1	200638635	200663378	+	1.84138 01	1.6593 088	3.4227 7	4.249 6977	2.985 226	3.536 2139
ENSG00000267519	1.62	MIR24-2	chr19	13945330	13947173	-	4.72304 44	4.5993 896	8.6144 75	9.827 661	7.844 5435	10.23 9091
ENSG00000254681	1.61	PKD1P5	chr16	18458147	18495797	-	1.31439 34	1.0079 812	1.4375 087	1.722 2397	2.123 2183	2.183 325
ENSG00000197989	1.61	SNHG12	chr1	28905050	28909495	-	3.87161 02	2.7489 998	4.3954 32	5.069 179	5.840 756	6.601 1386
ENSG00000247315	1.61	AL034548.1	chr20	278214	280961	+	3.02354 24	2.0685 76	2.8194 673	4.320 936	4.353 7407	3.915 0994
ENSG00000251260	1.61	WDFY3-AS1	chr4	85724411	85731544	+	1.00389 2	1.4533 712	2.2153 144	3.214 5927	1.323 8585	3.147 409
ENSG00000250182	1.60	EEF1A1P13	chr5	14652047	14653438	-	9.46529	6.2075 24	6.9742 58	12.66 7569	11.56 31275	11.53 3177
ENSG00000251143	1.60	RP11-849H4.4	chr11	71725337	71731956	+	3.97319 3	6.8831 76	5.2716 713	9.131 103	7.891 925	8.129 692
ENSG00000272888	1.59	AC013394.2	chr15	93425937	93441975	+	0.84985 04	1.4388 658	1.8384 619	2.750 0532	2.498 7977	1.538 8129
ENSG00000173295	1.58	FAM86B3P	chr8	8086117	8102387	+	0.99347 425	0.8375 4015	1.0657 19	1.652 3628	1.481 0578	1.727 7136
ENSG00000271581	1.58	XXbac- BPG248L24.12	chr6	31324424	31325414	+	10.9642 39	42.648 693	30.244 503	26.87 288	45.12 537	46.04 9137
ENSG00000234694	1.58	RP1-92O14.3	chr1	43820355	43824329	-	1.44880 64	2.5090 265	1.8279 302	3.333 9834	4.057 3297	1.937 6101
ENSG00000231970	1.58	RP11- 452K12.7	chr10	99160872	99179281	+	4.40716 4	1.8063 688	1.6439 916	2.576 6385	4.841 898	4.122 6187
ENSG00000258824	1.58	CTD- 2555O16.2	chr14	64889653	64915275	-	0.54871 875	0.3577 88	0.8857 678	2.461 6747	1.143 732	1.395 1521
ENSG00000224886	1.58	BEND3P3	chr10	81444250	81444739	-	0.14750 265	1.2424 577	0.5076 731	1.309 9444	0.603 35094	3.709 9042
ENSG00000259623	1.57	RP11-156E6.1	chr17	40004770	40007699	-	1.19021 55	0.7809 776	1.3159 648	2.417 0635	1.441 4529	1.755 5282

ENSG00000214184	1.57	AC012487.2	chr2	109123971	109150652	-	1.13318 16	1.5004 897	1.6879 808	2.188 2014	1.909 4998	2.645 182
ENSG00000231864	1.57	RP11- 229P13.23	chr9	139957987	139959033	+	2.14776 64	0.5192 181	1.6303 065	3.468 3156	2.224 2556	1.746 7672
ENSG00000259948	1.56	RP11- 326A19.5	chr15	89744322	89744999	-	0	0	0	0	2.250 7296	1.690 6314
ENSG00000203875	1.56	SNHG5	chr6	86370710	86388451	-	102.081 635	49.028 21	72.932 594	136.5 7454	115.0 8786	87.77 891
ENSG00000262944	1.56	RP11-473I1.9	chr16	9198705	9200618	+	4.37061 83	3.5058 978	5.2434 583	9.294 138	6.550 2	4.981 1077
ENSG00000223361	1.56	FTH1P10	chr5	17353804	17354733	-	1.12084 73	1.9096 965	3.1168 904	3.494 8602	3.568 4345	2.014 9157
ENSG00000234797	1.55	RPS3AP6	chr15	60060543	60061347	+	8.03755 8	4.6485 057	6.0571 284	10.87 507	9.480 915	8.230 892
ENSG00000186831	1.55	KRT17P2	chr17	18330175	18335162	+	2.67693 57	5.9414 525	8.2313 9	8.073 757	7.960 8884	7.615 3316
ENSG00000235298	1.55	RP11-575L7.8	chr9	86587148	86590692	+	0.99603 415	0.1577 4745	1.2684 056	1.084 6498	1.106 6335	3.923 4712
ENSG00000244398	1.55	RP11- 466H18.1	chr11	16996240	16996560	-	210.415 99	118.90 0734	159.80 867	323.5 744	266.5 1407	171.5 4987
ENSG00000232956	1.54	SNHG15	chr7	45022622	45026560	-	4.96357 63	4.2213 054	6.0738 254	7.899 4136	7.568 8367	7.776 898
ENSG00000265688	1.54	MAFG-AS1	chr17	79885705	79888629	+	2.63169 65	1.0524 323	1.6549 61	2.523 4575	3.253 9558	2.027 5886
ENSG00000205746	1.54	RP11- 1212A22.1	chr16	18428257	18488396	-	2.70093 97	1.9503 261	2.9362 893	3.497 657	3.365 603	4.758 7376
ENSG00000205763	1.53	RP9P	chr7	32956427	32982788	-	1.58876 7	2.6736	4.2468 457	3.486 325	3.794 6262	4.901 2694
ENSG00000260107	1.53	AC005606.15	chr16	2047655	2048375	+	9.51927 2	5.6106 744	7.2765 503	8.348 566	12.49 4355	13.30 5827
ENSG00000232445	1.53	RP11-132A1.4	chr7	100951627	100954266	+	6.63814 2	1.0719 961	3.6041 183	4.977 168	4.163 5494	4.395 196
ENSG00000204194	1.52	RPL12P1	chr6	33367836	33368333	-	8.81644 3	4.4001 51	10.382 515	8.702 698	12.67 3401	12.93 3394
ENSG00000254473	1.52	RP11-522I20.3	chr9	86322509	86328293	+	0.62335 21	0.6170 2996	0.9024 317	2.064 1985	1.439 8504	1.178 33
ENSG00000261971	1.52	RP11- 473M20.7	chr16	3101992	3109371	-	0.38817 36	0.9558 703	1.3753 506	1.559 6504	1.193 0581	2.581 4805
ENSG00000235725	1.51	AC007389.3	chr2	65816700	65867311	-	0.91346 866	0.7893 489	1.0001 628	1.792 3195	1.434 173	1.351 0778
ENSG00000229212	1.51	RP11-561C5.4	chr15	85747141	85778026	-	1.54617 49	2.1187 642	2.2953 222	2.896 4856	2.832 387	3.181 3629
ENSG00000259328	1.51	RP11- 152F13.7	chr15	82944773	82974312	+	1.00161 61	1.3405 782	1.1754 336	1.238 5976	1.825 4333	2.403 051
ENSG00000223396	1.51	RPS10P7	chr1	201487831	201499602	+	3.12924 72	1.9355 738	2.9419 267	4.337 2602	4.088 839	3.454 8473
ENSG00000269032	1.51	AC016629.7	chr19	59097525	59097723	+	0.32712 79	0.8532 135	0.6416 285	0	2.120 2471	1.611 3639
ENSG00000203394	1.51	RP5-930J4.4	chr1	21069480	21070455	+	4.33596 37	2.2852 726	2.7163 084	5.946 6834	4.018 1866	3.844 3263
ENSG00000225151	1.50	AC103965.1	chr15	84867600	84898888	-	7.19354 3	9.0674 305	8.6491 41	13.35 8168	12.31 3677	11.67 2823
ENSG00000244257	1.50	PKD1P1	chr16	16404198	16428047	+	2.76064 6	1.8415 892	2.8608 568	3.244 7884	3.438 1177	4.434 0997
ENSG00000267580	1.50	CTD- 2540B15.11	chr19	33790840	33792074	+	1.09239 59	2.0792 82	1.7315 762	2.528 3904	2.831 5682	1.866 6612
ENSG00000267151	1.50	MIR2117	chr17	41522075	41528568	+	1.46409 5	1.9914 546	1.7382 286	2.071 921	3.630 837	2.286 1967
ENSG00000270055	1.50	CTD- 3092A11.2	chr15	30780166	30782516	+	1.17078 79	0.9375 974	1.4183 607	2.055 2423	1.595 0989	1.714 8618
ENSG00000228989	1.50	AC133528.2	chr2	242629829	242633704	+	1.82867 69	1.5681 707	0.6850 529	3.815 1476	2.544 251	0.348 43844
ENSG00000262791	1.50	RP11-	chr17	1629042	1641879	-	0.61560	0.1723	0.3134	1.525	0.680	2.214

		961A15.1					89	9359	515	6633	255	0858
ENSG00000248092	1.50	NNT-AS1	chr5	43571696	43603332	-	2.29409 2	2.9645 55	2.8210 022	4.673 5754	3.726 381	3.718 8199
ENSG00000270232	-1.51	RP11-475J5.7	chr5	99381632	99382078	-	14.4310 055	8.3597 29	7.1948 752	7.179 785	5.739 9325	6.123 285
ENSG00000255125	-1.51	RP11-685M7.5	chr11	10804860	10823172	+	2.00767 5	1.9517 535	4.4939 42	1.263 7898	0.861 9299	4.028 4314
ENSG00000265415	-1.52	CTD-2510F5.4	chr17	57280038	57281190	-	2.19017 36	1.3541 911	1.1740 087	0.709 3079	0.546 59224	0.912 0858
ENSG00000232358	-1.52	RP5-955M13.4	chr20	49615907	49626556	+	10.9595 83	13.578 551	15.932 661	5.993 6986	10.71 9287	10.58 0886
ENSG00000231298	-1.52	LINC00704	chr10	4692377	4720346	-	3.50088 14	0.8419 1686	0.5089 952	0.122 57986 5	0	0.204 42243
ENSG00000269834	-1.52	CTD-3018O17.3	chr19	52892095	52901019	-	0.42566 574	3.0581 512	1.1532 478	0.938 63595	0.609 15244	0.861 5535
ENSG00000228649	-1.53	AC005682.5	chr7	22893797	22901021	+	2.84680 3	1.1739 081	1.4650 004	1.375 8489	0.801 59485	0.919 04426
ENSG00000250234	-1.53	CTD-2024P10.1	chr5	34656517	34657355	-	2.84991 8	3.6838 043	2.7612 894	1.368 9165	1.759 3302	3.366 3306
ENSG00000260549	-1.53	MT1L	chr16	56651388	56652730	+	32.0605 55	7.0488 276	13.661 485	8.433 363	8.580 506	11.88 6822
ENSG00000269983	-1.53	RP11-497H16.9	chr5	69745463	69746180	+	2.89407 18	2.6753 237	1.4972 115	1.057 8682	2.014 671	1.510 466
ENSG00000232656	-1.54	IDI2-AS1	chr10	1068606	1090138	+	8.97962	2.5590 303	4.5844 545	4.037 7965	2.612 7467	2.731 6613
ENSG00000273132	-1.54	RP11-350J20.12	chr6	150173598	150174328	+	4.93389 4	3.8961 434	1.3101 063	0.214 81651	2.748 0006	2.503 8617
ENSG00000262413	-1.54	RP11-498C9.3	chr17	79825597	79826428	+	131.380 23	83.971 855	65.720 39	57.11 3407	68.61 178	50.39 401
ENSG00000233144	-1.55	RP11-537A6.9	chr10	75141191	75143254	+	1.73480 49	1.7642 145	2.0247 85	0.579 00864	1.376 7893	1.206 9995
ENSG00000226963	-1.55	AC078883.4	chr2	173292502	173293331	-	3.65149	3.3694 608	3.6925 917	2.460 7828	1.833 3695	2.698 5145
ENSG00000234327	-1.55	AC012146.7	chr17	5014763	5017674	+	1.95390 83	2.4199 314	1.5286 9	0.984 7564	1.408 148	1.369 486
ENSG00000259153	-1.56	RP6-65G23.3	chr14	71276922	71282120	+	1.99365 48	2.7267 768	1.2656 376	1.045 882	0.530 8814	1.731 0688
ENSG00000258791	-1.58	LINC00520	chr14	56247864	56263406	-	0.64874 68	2.9121 616	2.3695 24	0.322 98782	1.232 0567	1.418 7404
ENSG00000267776	-1.58	AC006116.24	chr19	56888073	56888653	-	2.15047 86	3.7224 712	2.8212 99	1.832 4789	1.608 6241	1.933 2796
ENSG00000257181	-1.58	RP11-611O2.5	chr12	69235068	69237017	-	5.82705 8	7.4925 43	6.0171 41	4.169 096	3.725 6134	4.259 7303
ENSG00000267774	-1.58	RP11-2N1.2	chr18	57363691	57364644	+	4.9585	2.7897 93	5.2442 584	1.467 508	3.439 1992	3.611 8386
ENSG00000241749	-1.59	RPSAP52	chr12	66151800	66220754	-	3.20631 31	1.0826 937	1.4755 359	0.093 51292	0.572 23415	1.282 0605
ENSG00000268854	-1.59	CTD-2545M3.2	chr19	50983376	50986608	-	4.72710 04	2.3693 736	1.3470 087	3.467 783	1.066 7965	1.013 8669
ENSG00000224818	-1.59	RP11-134G8.8	chr1	201433511	201434274	-	3.14244 77	7.4775 295	3.8924 434	2.826 5297	2.930 994	2.728 9903
ENSG00000232815	-1.59	LINC00537	chr9	68413482	68414196	-	3.58854 34	3.3276 02	2.7501 92	2.034 8413	2.185 565	1.820 1643
ENSG00000241494	-1.61	RP11-796G6.1	chr14	102144280	102144717	+	3.58287 45	1.8213 323	2.3123 465	1.848 1115	1.181 2179	1.666 4281
ENSG00000258017	-1.62	RP11-386G11.10	chr12	49521565	49541652	+	27.8705 02	18.402 964	17.180 283	12.05 1238	12.65 8115	13.56 6694
ENSG0000034063	-1.62	UHRF1	chr19	4903092	4962165	+	6.74429 46	1.2855 517	2.7702 923	2.075 5224	1.767 3001	1.536 0543
ENSG00000254258	-1.62	RP11-398H6.1	chr8	140472305	140475259	-	4.54776 14	7.7839 007	6.3626 404	3.925 3066	3.584 0328	3.740 049

ENSG00000213412	-1.64	HNRNPA1P33	chr10	47133338	47133898	-	14.4520 4	8.3262 005	6.1332 52	8.466 77	4.197 107	4.750 102
ENSG00000126005	-1.64	MMP24-AS1	chr20	33804265	33865934	-	9.15239 3	7.7007 318	6.1065 493	5.459 143	5.335 1436	3.377 0957
ENSG00000172965	-1.65	MIR4435-1HG	chr2	111953927	112252677	-	3.55057 67	3.3823 805	3.6901 011	1.780 8807	2.279 2993	2.452 1575
ENSG00000253161	-1.65	RP11- 150O12.1	chr8	37278859	37411701	-	11.0290 21	12.441 466	9.9816 79	6.311 8887	8.075 069	6.012 6433
ENSG00000261068	-1.65	RP11-7K24.3	chr6	42059976	42061997	-	1.63533 17	8.6070 25	3.6644 359	1.894 0854	2.506 5024	2.411 0541
ENSG00000187653	-1.66	TMSB4XP8	chr4	91759652	91760263	-	8.07511 6	19.661 818	14.503 13	6.456 596	7.568 5825	10.31 9572
ENSG00000226005	-1.66	RP11- 464C19.3	chr10	3976711	3978005	+	1.07008 05	1.8116 003	2.9764 59	0.289 62108	0.678 89744	1.258 3201
ENSG00000250318	-1.67	CTA-963H5.5	chr22	31049864	31050801	-	1.83993 59	2.2982 81	1.8083 909	0.734 8674	0.672 6246	1.651 6936
ENSG00000240225	-1.67	ZNF542	chr19	56879468	56891197	+	2.40770 34	2.6245 768	2.5352 726	1.434 9543	1.558 2225	1.541 5956
ENSG00000240875	-1.68	LINC00886	chr3	156465135	156534851	-	8.04174 8	2.8910 108	2.4221 4	2.591 8982	2.259 8019	2.029 531
ENSG00000256663	-1.72	RP11- 424C20.2	chr12	20704524	20705946	+	4.26013 66	0.7228 8775	1.1924 789	0.947 2472	0.934 9184	0.820 1998
ENSG00000241990	-1.72	RP6-109B7.3	chr22	46449585	46453090	+	7.39844 27	3.9867 473	4.8413 033	3.091 525	3.469 6312	2.616 5955
ENSG00000245112	-1.74	SMARCA5-AS1	chr4	144434625	144435788	-	0.00956 0357	5.2469 225	0	0	0.159 80777	0
ENSG00000259330	-1.74	LINC00984	chr15	40617417	40618916	+	4.89751	3.6402 433	3.2306 47	2.593 4253	2.204 161	1.912 7241
ENSG00000235385	-1.75	GS1-600G8.5	chrX	13284167	13321571	-	4.73486 38	22.813 477	19.080 942	3.339 2224	14.33 58755	8.066 61
ENSG00000261065	-1.75	RP11-74C13.4	chr1	204100190	204101094	+	1.99908 89	2.6730 783	0.4170 3206	0.414 32646	0.868 4934	0
ENSG00000253618	-1.76	GRPEL2-AS1	chr5	148727679	148737205	-	1.68886 54	3.2317 665	0	0.215 5406	0.803 7886	0
ENSG00000258232	-1.77	RP11- 161H23.5	chr12	49658939	49667089	-	33.3097 57	28.819 185	28.912 59	14.44 1423	19.74 7879	17.62 5807
ENSG00000272734	-1.79	ADIRF-AS1	chr10	88725102	88731068	-	5.43516 25	12.950 257	11.457 844	4.700 653	5.038 381	5.980 5155
ENSG00000234546	-1.81	RP3-510D11.2	chr1	9242263	9252148	+	11.2780 56	4.5766 697	5.7239 7	2.883 4605	5.230 489	3.284 4615
ENSG00000258908	-1.82	RP11- 203M5.8	chr14	20942948	20945248	-	7.68502 6	3.8135 834	4.4676 06	2.660 3262	2.985 3904	2.748 8363
ENSG00000272159	-1.83	RP11- 350N15.6	chr8	38265566	38266260	-	7.30503 03	2.4662 013	2.9248 87	1.585 6465	2.061 9986	2.634 9046
ENSG00000226721	-1.84	EEF1DP2	chr9	95599108	95599950	-	0	6.2755 63	0	0	0	0
ENSG00000204556	-1.86	CTD-2514C3.1	chr20	25999468	26001320	+	3.08402 75	2.3788 395	1.0570 529	1.181 9186	0.448 66753	1.024 0932
ENSG00000197358	-1.86	BNIP3P1	chr14	28733596	28735180	+	4.02136 66	6.1853 94	4.5065 994	2.335 3097	2.544 7326	2.919 4083
ENSG00000235237	-1.88	RP1-151B14.6	chr22	37562797	37578890	+	5.69870 1	5.5882 27	3.9706 326	2.796 2503	2.296 2286	2.966 6011
ENSG00000237499	-1.89	RP11-356I2.4	chr6	138144810	138189370	-	2.07288 89	2.8674 767	2.2776 296	0.910 7068	1.268 385	1.577 585
ENSG00000235847	-1.89	LDHAP7	chr2	85004383	85005347	-	2.74156 7	2.0703 194	1.2152 584	0.508 1032	1.017 5685	0.249 74425
ENSG00000273038	-1.90	RP11- 479G22.8	chr10	33176189	33178239	-	2.26414 06	2.7482 634	1.5507 42	0.526 8092	1.410 5644	0.987 04404
ENSG00000231991	-1.91	ANXA2P2	chr9	33624223	33625532	+	134.588 53	110.10 068	128.57 472	52.56 764	66.60 7285	77.69 0636
ENSG00000256929	-1.91	AC067852.1	chr17	40673736	40674065	+	3.25574	8.0733	4.7597	3.042	3.687	1.590
ENSG00000255959	-1.92	RP11-	chr11	60603469	60610438	-	5.55156	2.3799	2.5062	1.125	1.753	2.367

		804A23.2					56	477	263	2497	554	571
ENSG00000251292	-1.93	RP11-380P13.2	chr4	23724885	23735202	-	38.60298	68.663376	43.137657	25.190517	34.787884	18.087368
ENSG00000262410	-1.96	RP11-388C12.8	chr17	80702944	80703585	+	2.0405107	3.6950738	4.62204	2.3328707	0.49339446	1.9949535
ENSG00000266088	-1.97	RP5-1028K7.2	chr17	38673278	38683254	+	4.3136535	1.3317922	1.334867	0.23217432	0.22915247	0.5862707
ENSG00000255234	-1.98	RP11-727A23.10	chr11	82997171	83134559	+	74.264435	131.67998	93.62611	40.383236	64.519875	45.11061
ENSG00000231826	-1.98	AC016735.2	chr2	43254992	43266686	-	16.923	4.0733547	5.240191	4.893341	3.788711	2.493266
ENSG00000235823	-2.00	LINC00263	chr10	102133372	102143125	+	2.904085	3.2670798	2.4657273	1.211951	1.8418641	1.3118396
ENSG00000265992	-2.00	ESRG	chr3	54666149	54673884	-	233.58145	439.6651	314.1879	138.61157	202.88977	142.82372
ENSG00000260834	-2.02	RP11-256I9.2	chr16	65224876	65268817	-	56.87727	96.663605	70.33388	29.448397	46.84315	33.758797
ENSG00000256940	-2.06	RP11-783K16.5	chr11	64013436	64015689	+	13.28241	9.2834835	8.433685	6.031071	4.4094415	4.4594874
ENSG00000273387	-2.07	RP3-412A9.16	chr22	31478142	31479551	-	3.5070229	1.896202	1.4603726	1.1005665	0.39510265	0.5181117
ENSG00000255021	-2.08	RP11-536I6.2	chr3	14313873	14345345	-	14.229764	26.088594	15.77334	7.2628064	12.217939	7.3267016
ENSG00000100181	-2.08	TPTEP1	chr22	17082777	17179632	+	20.285084	40.96005	27.09847	11.958863	17.235292	12.10646
ENSG00000236673	-2.09	RP11-69I8.2	chr6	132223103	132241705	+	8.285399	16.666319	10.295293	4.810559	6.5929637	4.929272
ENSG00000226792	-2.09	LINC00371	chr13	51656984	51746524	-	9.052619	16.953245	10.620668	5.2847075	6.8860707	4.8756647
ENSG00000236824	-2.12	BCYRN1	chr2	47558199	47571656	+	12.523726	24.176477	18.561161	5.1182065	7.3273034	15.67111
ENSG00000224081	-2.12	LINC01057	chr1	95104017	95285837	-	59.76518	112.70979	71.806335	32.73587	50.640713	30.50008
ENSG00000225383	-2.15	SFTA1P	chr10	10826400	10836943	-	33.883236	11.625587	16.437605	5.999579	6.5704207	16.636505
ENSG00000219445	-2.15	RP11-3B12.3	chr7	124869633	124904345	-	5.3039045	10.288663	7.675313	3.506465	4.1590476	2.9025939
ENSG00000233218	-2.25	SNX18P16	chr1	142688379	142688925	-	1.7177182	2.6481433	3.362058	0.43026203	1.3511974	0.33988443
ENSG00000233461	-2.25	RP11-295G20.2	chr1	231658134	231664302	-	4.3447165	5.260303	4.8380613	0.9010945	2.7094634	3.6013868
ENSG00000269495	-2.34	CTB-147C22.8	chr19	51453441	51466905	+	4.9118886	14.511923	12.695248	1.7244989	4.9518948	8.315971
ENSG00000257219	-2.41	RP11-54A9.1	chr12	76653619	76698911	+	3.866573	2.7792904	2.3597717	0.30285415	0.94948655	1.8060194
ENSG00000230439	-2.41	RP11-488P3.1	chr1	94218480	94241000	-	16.71741	2.3367712	4.880771	2.3431618	2.6412401	2.188129
ENSG00000214110	-2.44	LDHAP4	chr9	14921335	14922332	-	6.771377	6.329879	5.676562	2.1420772	2.5751235	3.043416
ENSG00000267279	-2.45	RP11-879F14.2	chr18	59252979	59274149	-	6.177033	1.2901783	1.8376158	0.89223325	0.91465175	0.94758815
ENSG00000254740	-2.63	RP11-334E6.3	chr11	119243416	119252323	+	1.6593597	5.261143	5.3788548	1.0410688	1.6390346	1.5145406
ENSG00000268621	-2.68	AC006262.5	chr19	46692423	46706340	-	2.935837	5.8567724	3.714047	1.0019716	2.6371474	1.2505352
ENSG00000271573	-2.69	RP11-96C23.12	chr10	88786941	88787234	+	1.966702	9.996787	7.010724	0.4366481	2.154825	3.302377
ENSG00000257042	-2.69	RP11-993B23.3	chr12	28111450	28122746	+	25.18919	26.20775	21.461538	4.8206534	9.942112	15.1163
ENSG00000255198	-2.71	SNHG9	chr16	2014960	2015510	+	5.082953	5.41024	2.8378108	1.3202466	1.1497614	2.5868635
ENSG00000261780	-2.74	CTD-2354A18.1	chr18	70992176	71017113	+	7.452665	13.35417	8.380511	2.9860015	4.8623896	2.8057775

ENSG00000258177	-2.75	RP11-394J1.2	chr12	96616575	96617751	-	11.3985 49	7.2211 123	5.6029 59	1.443 8808	3.685 8037	4.163 6434
ENSG00000255729	-2.78	AC005618.1	chr5	140699661	140700339	-	3.83701 63	7.4735 03	7.0763 206	0.756 8286	4.717 6375	1.997 0802
ENSG00000238266	-2.79	LINC00707	chr10	6821560	6884868	+	3.58438 56	4.5337 21	3.8257 384	1.909 5888	1.506 412	0.787 2698
ENSG00000239492	-2.98	FAM25HP	chr10	47740330	47747200	+	15.1145 64	31.412 348	26.541 6	5.972 542	9.727 91	8.178 021
ENSG00000253746	-3.17	RP11- 527N22.2	chr8	37262957	37264242	-	3.75136 5	4.0315 757	2.6437 647	0.703 32074	0.949 91237	1.250 8819
ENSG00000256288	-3.22	RP11- 277P12.10	chr12	10485460	10490891	-	6.18079 3	7.4155 67	4.6269 43	0.939 5672	3.611 7387	1.750 629
ENSG00000242147	-4.40	RP13- 463N16.6	chr10	5636954	5638081	-	18.6134 9	1.4156 137	3.2376 003	0.521 8478	0.030 29739 3	0.474 20192
ENSG00000229647	-4.79	AC007879.7	chr2	208104374	208110611	+	31.3377 27	7.7880 12	10.499 597	3.835 9287	3.087 454	1.968 523

Table S2: Comparison of FAM83H-AS1 RPKM (Reads Per Kilobase Million) with other previously published functional LncRNAs

LncRNA	ENSEMBL ID	Avg. RPKM GFP	Avg. RPKM 16E6	Known function
FAM83H-AS1	ENSG00000203499	5.876310667	9.9992845	
XIST	ENSG00000229807	8.137879	12.21193233	miRNA sponge in cervical cancer (Zhu et al 2018)
HOTAIR	ENSG00000228630	1.817852513	1.5785181	miRNA sponge in cervical cancer (Liu et al 2018)
NEAT1	ENSG00000245532	34.23374033	27.70111733	miRNA sponge in cervical cancer (Han et al 2018)
TINCR	ENSG00000223573	16.26229083	16.07113433	miRNA sponge in gastric cancer (Chen et al 2017) and breast cancer (Liu et al 2018)
H19	ENSG00000130600	3.3026914	15.14256633	miRNA sponge in cervical cancer (Ou et al 2018)

Table S3: Primers used in this study for qRT-PCR.

Primers	Sequence (5'-3')
FAM83H-AS1 FWD FAM83H-AS1 REV	TCCAATAAACAGGGCAGAC CAAGATCACCACACCCCTCT
Gapdh FWD Gapdh REV	CCACTCCTCCACCTTTGAC ACCCTGTTGCTGTAGCCA
UBC FWD UBC REV	GATTTGGGTCGCAGTTCTTG CCTTATCTTGATCTTTCCTTG
Beta-Actin FWD Beta-Actin REV	AGCACAGAGCCTCGCCTTT CCACGATGGAGGGGAAGAC
U6 FWD U6 REV	GTGCTCGCTTCGGCAGCACATAT AAAAATATGGAACGCTTCACGAA
<i>C. Elegans</i> ama-1 FWD <i>C. Elegans</i> ama-1 REV	GGAGCTCGAGTGGATCTTCG GCGCAGAGAGTATCCTGGAC
HPV-16E6 FWD HPV-16E6 REV	AATGTTTCAGGACCCACAGG ACTGTTGCTTGCAGTACACACA
HPV-16E7 FWD HPV-16E7 REV	ACAAGCAGAACCGGACAGAG GCCCATTAAACAGGTCTTCCA
p53 FWD p53 REV	TTTGGGTCTTTGAACCTTG CCACAACAAAACACCAGTGC
p300 FWD p300 REV	AGCGGCCTAAACTCTCATCTC CACCATTGGTTAGTCCCAATTC
FAM83H FWD FAM83H REV	CGACAAGTGCCGTGTCAACC ACTTCCCAGTGCGGCAGTAG
H19 FWD H19 REV	AGTGGACTTGGTGACGCTGTAT CTCCTGAGAGCTCATCACTCC
MIR205HG FWD MIR205HG REV	TCATTAAAGAGAGAAATCAACTATTCA TAAAGCACCCAGTTCAGCAG
HOXC-AS5 FWD HOXC-AS5 REV	CTCCTTGGAGCAGTACACCTG GTCCATTTGCTCTCCACTTGA
GAS5 FWD GAS5 REV	CTTCTGGGCTCAAGTGATCCT TTGTGCCATGAGACTCCATCAG
LINC00963 FWD LINC00963 REV	AACTGCCTTTGGAAGCAAGTAG ACTGCTCTAGTCCAGGAGTTCG
SNHG15 FWD SNHG15 REV	GCTGAGGTGACGGTCTCAA GCCTCCCAGTTTCATGGACA
MAFG-AS1 FWD MAFG-AS1 REV	AGGACTCGGGAGGAAGATAAAC AACACCTTCAGCTCTCTGCTG
RP6-65G23.3 FWD RP6-65G23.3 REV	GAGATAGGAGGCCATAATGTTTC CTGCAACCAGGTGGAAGTCA
GS1-600G8.5 FWD	ACTGTCCGGTTCCTGTGGG

GS1-600G8.5 REV	CGGTGCCACTGGGTCATTTT
RP3-510D11.2 FWD	CCAGACCGACGGGACAGCG
RP3-510D11.2 REV	GCTTCCCTGTCCTCCTCCTA
RP11-479G22.8 FWD	TGGGAAACTAAAACCATTTAAGC
RP11-479G22.8 REV	GCTCAGAACATGGTTTTCTCTG
LINC01057 FWD	GTGAATTCCTTGAAGATGAGG
LINC01057 REV	TCATTCACTCAACCACTGAACC
SFTA1P FWD	CAGCATTCCAGGTGGGCTTT
SFTA1P REV	CCTTGTTTGGCTTACTCGTGC
RP13-463N16.6 FWD	TTGGAAATCACTCCTTCCACTT
RP13-463N16.6 REV	CAGAGACAATTCAACTCCCACA
AC007879.7 FWD	ACAGGGAGCCAGGACACC
AC007879.7 REV	GGAACCAGCACCAGGAACC

XIII. Supplementary Methods

A. Cell Culture

Primary human foreskin keratinocytes (HEKa) were cultured as described by supplier (EpiLife® Medium supplemented with Human Keratinocyte Growth Supplement (HKGS)). Primary human cervical keratinocytes (HCK) and J2-3T3 mouse fibroblast feeders were cultured as described by Alison McBride's laboratory (NIH, Bethesda, MD)¹. HCK were maintained in F-media [3:1 F12:DMEM with 5% Fetal Bovine Serum (FBS), 0.4ug/mL Hydrocortisone, 5ug/mL Insulin, 8.4ng/mL Cholera Toxin, 10ng/mL Epidermal Growth Factor, 24ug/mL Adenine, 10U/mL Penicillin, 10ug/mL Streptomycin, 2mM L-Glutamine, Amphotericin B] containing 10uM Y-27632 ROCK Inhibitor (Tocris). HCK cells were co-cultured with J2-3T3 feeder cells rendered mitotically inactive in 8ug/mL Mitomycin C for 3 hours; these growth-arrested feeder cells were replenished every 3-4 days. J2-3T3 stock cells were cultured in Feeder media [Dulbecco's Modified Eagles Medium (DMEM), high glucose, 10% Newborn Calf Serum, 2mM L-Glutamine, 10U/mL Penicillin, 10ug/mL Streptomycin, Amphotericin B]. J2-3T3 feeder cells are sensitive to trypsin; HCK, J2-3T3 stock, and growth-arrested J2-3T3 cells were centrifuged to remove the trypsin when passaged. HPV-16 positive CaSki cells were cultured in Roswell Park Memorial Institute (RPMI)-1640 medium containing 10% FBS, Streptomycin-Penicillin, HEPES Buffer, Amphotericin B. HPV-16 positive W12/20863 and W12/201402 and HPV-31b positive cervical (CIN-612) were co-cultured in E-media (1:1 DMEM:Ham's Nutrient Mix F12 medium supplemented with 5% FBS, 10mM HEPES Buffer, Penicillin-Streptomycin, Amphotericin B, 0.02μM Triiodothyronine, 0.4μg/mL Hydrocortisone, 0.1μg/mL Cholera Toxin, 5μg/mL Transferrin, 180μM Adenine, 5μg/mL Insulin) with murine 3T3 fibroblast cells (3T3M) that were rendered mitotically inactive with 4ug/mL Mitomycin C for at least 2 hours. HPV-negative cervical C33A cells, HPV-18 positive cervical (HeLa), HPV-16 positive HNSCC cell lines UMSCC-47 and -104, HPV negative HNSCC cell line UMSCC-1, HEK293T, and 3T3M cells were all cultured in DMEM supplemented with 10% FBS, Penicillin-Streptomycin, HEPES Buffer, and Amphotericin B. To passage stock cells, 0.25% trypsin was utilized. Growth-arrested J2-3T3 and 3T3M fibroblasts were removed with Versene solution and gently pipetting, prior to pelleting and when otherwise desired. Cells were cultured at 37°C in a humidified 5% CO₂ cell culture incubator.

B. Generation of stable cell lines

Plasmids MSCV-N-GFP (Plasmid #37855), MSCV-N-16E6 (Plasmid #37875), and MSCV-N-16E7 (Plasmid #37881) were purchased from Addgene. For retroviral production, these plasmids were

co-transfected with packaging (pCL-ECO) and envelope (pVSV-GF) plasmids into HEK293T cells using calcium phosphate co-precipitation. 12-16 hours later, HEK293T media was replaced with fresh target cell media. After 48-hour incubation, retrovirus-containing target cell media was harvested, filtered (0.45µm), and mixed with polybrene (4µg/mL final concentration) to increase infection efficiency. 24 hours before infection, co-cultured J2-3T3 fibroblasts were removed with Versene solution and gentle pipetting from HCK cells. HEKa and HCK cells were infected with recombinant retroviruses for 8 hours, virus was removed and replaced with fresh target cell media. Growth arrested J2-3T3 were added to infected HCK and replenished every 3 days. Cells were allowed to recover for 24 hours prior to 72-hour puromycin selection (HEKa-2.5µg/mL; HCK-3µg/mL). Cells were allowed to recover and expression of exogenous HPV-16 E6 and E7, as desired, was confirmed by RT-PCR (Fig. S1, Fig. S3).

C. Generation of HPV-16 positive cervical JAMM-16 cell line

According to Buck et al. protocol described previously², the HPV-16 insert was cut out of pBR322 HPV-16 plasmid by performing a restriction digestion of two separate 25ug plasmid reactions in 225uL each with the restriction enzyme BamHI (NEB) according to manufacturer's recommended conditions (37°C for 2 hours). Digested samples were then PCR purified by QIAquick® PCR Purification Kit (Qiagen) resuspended in 200uL Buffer TE. Purified samples were then ligated under dilute conditions (9mL total volume) by adding 1X Ligase Buffer and 6uL of high concentration (2m U/mL) T4 DNA Ligase (NEB) and incubated at 16°C overnight. <25ug in 9mL ligation reactions were conducted to avoid concatemer formation. The ligated samples were then treated with 4.5mL of 7.5M Ammonium acetate and mixed. 35mL of 95% ethanol was added to each, tubes were mixed, and incubated at 4°C overnight. The next day, the samples were brought back to room temperature and centrifuged at ~5,000 x g at room temperature for 1 hour. Pellets were washed with 70% ethanol, spun briefly, and washed again with 70% ethanol. The pellets were spun one last time to remove any residual ethanol, then air dried for several minutes. Pellets were resuspended in 100uL Buffer TE each, pellets were combined, and Nanodrop was used to calculate concentration retrieved. The entire product was run in 1% agarose gel for 3 hours. Desired band [expect to see supercoiled and relaxed circular bacterial backbone (~1.8 and ~3.2 kb) and supercoiled and relaxed circular (nicked) HPV genome (~6.2 and >16 kb)] was cut out and gel extracted with QIAquick® Gel Extraction Kit (Qiagen). Before transfection of the HPV-16 genomes, growth-arrested J2-3T3 fibroblasts were removed by Versene solution and pipetting from co-cultured primary normal cervical keratinocytes (HCK). HCK were plated in a 6-well plate on day prior to transfecting according to manufacturer's protocol for Lipofectamine LTX (Invitrogen). On day of transfection, HCK were at about 70% confluency. Media was replenished on each well. For each well, 2.8ug of re-ligated

HPV-16 insert was diluted in 150uL Opti-MEM I Reduced Serum Media with 2.5uL PLUS Reagent, 15uL Lipofectamine LTX was diluted in Opti-MEM I Reduced Serum Media. Diluted DNA and diluted Lipofectamine LTX/Plus complex were combined and incubated for 5 minutes at room temperature prior to adding to well. The transfected cells were passaged at least 10-15 times so HPV-16 immortalized cells could overgrowth the untransfected HCK cells prior to experimentation.

D. RNA Extraction

Total RNA of cultured cell lines and human tissues was purified with TRIzol reagent (Invitrogen) and treated with Turbo DNA-free DNase (20 minutes at 37°C, Ambion) according to manufacturers' protocols. RNA concentration was determined with a Nanodrop 2000 Spectrophotometer (Thermo Scientific).

E. cDNA synthesis, qualitative or quantitative RT-PCR

RNA was reverse transcribed to complementary DNA (cDNA) under standard iScript cDNA synthesis kit (Bio-Rad) instructions in T100™ Thermal Cycler (Bio-Rad). Exogenous spike of *C. Elegans* (70ng) was added for normalization of results if GAPDH mRNA levels were altered during processing (cellular fractionation). RT-PCR or qRT-PCR using SsoAdvanced™ Universal SYBR® Green enzyme was performed according to manufacturer's protocol (Bio-Rad) in CFX Connect™ Real-Time System (Bio-Rad). The oligonucleotide primer sequences can be found in Supplementary Table S3. GAPDH was primarily used as housekeeping gene, however Ubiquitin C (UBC) or Glucuronidase Beta (GUSB) were used if GAPDH mRNA levels were altered with significant apoptosis (siRNA FAM83H-AS1 vs. siRNA-CTRL) and ama-1 was used for cellular fractionation studies. To validate cell fractionation efficiency, U6 snRNA was used as nuclear control and mature mRNA beta-actin as cytoplasmic control. Qualitative RT-PCR product was resolved by agarose gel electrophoresis to visualize or quantitative RT-PCR (qRT-PCR) relative expression was calculated using the double delta CT method (relative expression = $2^{-\Delta\Delta CT}$; where $\Delta CT = CT(\text{Target RNA}) - CT(\text{mRNA endogenous GAPDH/UBC control or } C. \text{ Elegans mRNA Ama-1 was exogenous spike control})$). Fold changes were calculated relative to control siRNA or the mean value of normal samples.

F. Western Blot analysis

Cell lysates were boiled in sample buffer, separated by SDS-PAGE, and transferred onto Immobilon-Fl polyvinylidene difluoride membranes. After blocking for 1 hour with 5% milk-TBST buffer (5% non-fat dry milk, 25mM Tris-HCl pH 8.0, 125mM NaCl, 0.5% Tween-20), the membranes were incubated with primary antibodies against human p53 (Cell Signaling, #9282) or Actin (C-11) (Santa Cruz, sc-1615) overnight at 4°C. The membranes were then washed with TBST, incubated with appropriate horseradish peroxidase (HRP)-conjugated secondary antibodies [anti-goat IgG (Santa Cruz, sc-2020); anti-mouse IgG (Thermo Scientific)] for 1 hour at room temperature, washed with TBST, and the proteins were detected on the membrane using Pierce SuperSignal West Pico or Femto Maximum Sensitivity Substrate (Thermo Scientific) chemiluminescence.

G. Transient transfection

The following siRNAs were used to knock down FAM83H-AS1: Lincode Human FAM83H-AS1 siRNA - Set of 4 (Dharmacon, RU-188909-00-0002), Lincode Human FAM83H-AS1 siRNA SMARTpool (Dharmacon, R-188909-00-000), Lincode Non-targeting siRNA #1 (Dharmacon, D-001320-01-05). Two different siRNA against HPV-16E6: 5'-GAGGUAUAUGACUUUGCUUTT-3' (Dharmacon) and 5'-UCCAUAUGCUGUAUGUGAUTT-3' (siRNA 209, Dharmacon), two different siRNA against p53 (NEB, #2011S; Thermo Scientific Dharmacon®, J-003329-16), two different ONTARGETplus Human siRNA against p300 (Dharmacon, J-003486-11-0002; Dharmacon, J-003486-12-0002), and CTLsiRNA (Ambion, AM16104) were used.

HCK, CaSki, and W12/201402 cells were transiently transfected using standard Lipofectamine® RNAiMAX protocol (siRNA FAM83H-AS1 and siRNA HPV-16E6 #1). W12/201402 and CaSki cells were transfected with reverse transfection protocol (siRNA HPV-16 E6 #2 to improve transfection efficiency. For standard transfection, cells were plated at appropriate confluency to be about 70% confluent for transfection 24 hours post-plating. Growth-arrested J2-3T3 fibroblasts were co-cultured with HCK, and on day of transfection removed with Versene solution and gently pipetting. On day of CaSki, W12/201402, and HCK standard transfection, media was replenished on each well. For each well, Opti-MEM I Reduced Serum Media was used to dilute Lipofectamine RNAiMAX (9uL per well of 6-well), as well as siRNA (40pmol per well of 6-well). Diluted RNAiMAX was combined with diluted siRNA, incubated 5 minutes at room temperature and added to each well. Cells were incubated 48 hours in media containing siRNA prior to harvesting, unless otherwise stated.

H. Cellular fractionation

Growth arrested 3T3M cells were removed by Versene solution and pipetting the day prior to fractionation for the W12/201402. Cellular fractionation protocol was modified from previous publication³. On ice, CaSki (~70% confluent) were washed two times with cold PBS, scraped, pelleted, and resuspended in RSB Buffer (10-1.25mM Tris, pH 7.4; 10-1.25mM NaCl; 3-0.38mM MgCl₂), incubated on ice for 5 minutes, centrifuged (1500rpm for 4 minutes at 4°C). RSB Buffer was removed and swollen pellets were resuspended in RSB-G40 Buffer [2.5-0.63mM Tris, pH 7.4; 2.5-0.63mM NaCl; 0.75-0.19mM MgCl₂; 2.5-0.63% glycerol; 0.125-0.03% NP-40; 0.125-0.03mM DTT; 100U/mL RNasin Plus RNase Inhibitor (Promega)] in ≥ 4 times volume of pellet. Cells were disrupted with Dounce homogenizer; homogenate was centrifuged (4000 rpm, 4 minutes) to pellet nuclear fraction. The cytoplasmic fraction (supernatant) was collected. The pelleted nuclear fraction was resuspended in RSB-G40 Buffer, 3.3% of sodium deoxycholate and 6.6% of Tween20 were added to the volume of RSB-G40 (10% of final volume was sodium deoxycholate and 10% was Tween 20), samples were tapped to mix, incubated on ice for 5 minutes, and pelleted by centrifugation (7000rpm, 3 minutes). The supernatant was collected and combined with first cytoplasmic fraction, and nuclear pellet was washed two times with RSB-G40 Buffer, and then resuspended in RSB-G40 Buffer. Fractionated cells were Turbo DNase I-treated, according to manufacturer's instructions (NEB), prior to RNA extraction to remove bound DNA and improve recovery of RNA. Total cell (second pellet), cytoplasmic, and nuclear RNA was extracted using TRIzol Reagent (Life Technologies) per manufacturer's instructions (Glycoblue was added during isopropanol precipitation step due to small amounts of RNA recovered from nuclear fraction) and treated with Turbo DNase (Ambion) for 20 minutes at 37°C. qRT-PCR analysis was performed using cDNA generated using equal concentration of RNA with exogenous *C. elegans* RNA (70ng) added for normalization.

I. Cell proliferation assay

Cell proliferation was monitored using Cell Counting Kit - 8 (CKK-8) (Sigma-Aldrich). 24 hours (CaSki) or 48 hours (201402) after transfection with FAM83H-AS1 siRNA, CaSki cells were detached with 0.05% trypsin and W12/201402 cells were detached with 0.25% trypsin. After cell detachment, trypsin was inactivated with trypsin neutralizer (CaSki) or E-media containing serum (201402), pelleted, resuspended in appropriate media containing serum, and plated at desired concentration (3E3 cells per well and of four separate 96-well plates; technical triplicates). In equal density per well, 3T3M fibroblasts treated with mitomycin C were added to the W12/201402 cells and fibroblast alone control wells were plated. At desired time (48, 72, or 96h post-plating), CKK-8 dye (100uL/well) was added to each well,

incubated for 2 hours (CaSki) or 3 hours (201402). Then, absorbance was read with a spectrophotometer (BioTek Synergy H1 Hybrid Reader) (OD₄₅₀) according to manufacturers' protocols. Absorbances of feeders alone were subtracted from W12/201402 co-cultured with feeders. Biological replicate was conducted.

J. Flow-cytometric analysis of cell cycle.

3T3M fibroblasts were stained with CellTrace™ Far Red Cell Proliferation Kit (Invitrogen) the day before plating transfected 201402 cells. 24 hours (CaSki) or 48 hours (201402) after transfection with FAM83H-AS1 siRNA, cells were detached with 0.05% trypsin (CaSki) or 0.25% trypsin (201402). After cell detachment, trypsin was inactivated with trypsin neutralizer (CaSki) or media containing serum (201402), and cells were plated at desired concentration (4E6 for CaSki and 2.3E6 for 201402) in 100mm dishes in media containing fetal bovine serum (FBS). Mitomycin C treated, far red stained 3T3M cells were added to the 201402 cells. Cells were allowed to attach (about 5 hours) then were washed 3 times with 1X D-PBS to remove residual FBS. Cells were then serum-starved for 24 hours. After incubation, cells were trypsinized with 0.25% trypsin, suspended in serum-free media (to avoid stimulation by serum), washed with 1X PBS, pelleted, and resuspended in PBS (200μL). Harvested cells were then fixed by adding 70% cold ethanol (1.8mL) while vortexing the cells and stored at 4°C until processing. Cells in ethanol were pelleted, washed with PBS, centrifuged, and resuspended in room temperature 0.2% Tween 20 in PBS (300-1,000μL, depending on pellet size). 100uL of each sample was placed in a U-bottom 96-well plate (3 wells/technical triplicates) and incubated 15 minutes at 37°C. PBS (100uL) was added to each well, cells were pelleted, resuspended in 10uL of RNase A-PBS (180ug/mL stock), tapped gently to mix, and incubated at room temperature for 15 minutes. After incubation, 20uL of PI-PBS (final concentration of 50μg/mL) was added to each well, pipetted to mix, and incubated for 15 minutes at room temperature. After incubation, the volume was brought up to 300uL and analyzed in the dark by flow cytometry (Fortessa S10). Gates were placed around 201402 population. Biological replicate was conducted.

K. Transwell migration assay

24 hours (CaSki) or 48 hours (201402) after transfection with FAM83H-AS1 siRNA, cells were detached with 0.05% trypsin. After cell detachment, trypsin was inactivated with trypsin neutralizer, pelleted, resuspended in serum-free media, and cells were plated [1E5 (CaSki) or 0.75E5 (201402), 200uL well, technical duplicates] in upper chambers of 24-well transwell (8mm pore size) with 800uL

serum-free media in lower chamber. In equal density per well, 3T3M fibroblasts treated with mitomycin C were added to the W12/201402 cells and fibroblast alone control wells were plated. Cells were then allowed to attach for ~3 hours prior to adding fetal bovine serum (final concentration of 20% for CaSki and 5% for W12/201402) to lower chamber. After 24 hours, media in upper chamber was replenished with appropriate fresh serum free media and lower chamber with appropriate fresh media containing with FBS chemoattractant. 48 hours after cells were plated, lower chamber/underside of transwell and upper chamber of transwell were washed with 25% D-PBS. D-PBS was removed and cells from the top of the upper chamber transwell membrane (non-migrated cells) were wiped away using a cotton swab. The lower chamber/underside of the transwell and upper chamber of transwell were washed again with 25% D-PBS and then the underside of the transwell (migrated cells) was fixed with 4% formaldehyde in 25% D-PBS in the lower chamber for 5 minutes at room temperature. The formaldehyde solution was removed and the lower chamber/underside of transwell and upper chamber of transwell were washed with 25% D-PBS. The cells on the underside of the transwell were then stained using 0.5% crystal violet in 20% ethanol in water. The stain was added to the lower chamber, submerging the transwell for 20 minutes at room temperature. The stain was removed and the lower chamber/underside of the transwell and upper chamber of transwell washed 3 times with 25% D-PBS. The top of the upper chamber transwell membrane was then cleaned with a cotton swab. The transwells were allowed to dry overnight and then the underside of the transwell imaged with an Olympus MVX10 microscope. Migration was quantitated with ImageJ. Feeders alone were subtracted from W12/201402 co-cultured with feeders. Biological replicate was conducted.

L. Flow-cytometric analysis of apoptosis

3T3M fibroblasts were stained with CellTrace™ Far Red Cell Proliferation Kit (Invitrogen) the day before plating transfected 201402 cells. 24 hours (CaSki) or 48 hours (201402) after transfection with FAM83H-AS1 siRNA, cells were detached with 0.05% trypsin (CaSki) or 0.25% trypsin (201402). After cell detachment, trypsin was inactivated with trypsin neutralizer (CaSki) or media containing serum (201402), pelleted, resuspended in appropriate media containing serum, and cells were plated in 100mm dishes at desired concentration (1.75E6 for CaSki and 1.1E6 for 201402). At 24, 48, and 72h later post-plating, attached and floating cells were collected. On day of harvest, media containing dead cells was collected, PBS was added to the plates, collected, and added to the media. Attached cells were trypsinized with 0.25% trypsin, cells were pelleted (2 to 5E6) at 4°C, resuspended in PBS, and 100uL of each sample was aliquotted into 3 wells of a 96 well plate (technical triplicates). Cells were pelleted, PBS was removed, Annexin V Binding Buffer (100uL; diluted with autoclaved milliQ water) was added to each

tube, Annexin V-FITC (5uL) was added to each tube, and incubated in the dark at room temperature for 10 minutes. After incubation, cold Binding buffer (200uL) was added to each tube, cells were pelleted, washed with cold Binding buffer two times, incubated on ice for 5 minutes in propidium iodine in Binding buffer (2ug/mL final concentration, 300uL), and analyzed by flow cytometry (Fortessa S10) immediately. Gates were placed around 201402 population. Biological replicate was conducted.

M. Human cervical tissue specimen

10 human specimens were obtained from patients under the tissue collection protocol [Prognostic Marker (IRB0406147)] at the University of Pittsburgh. Patient samples were categorized as cervical intraepithelial neoplasia stage 3 (CIN3), invasive cervical cancer (CaCx), or non-cancerous.

N. TCGA Analysis

The Cancer Genome Atlas (TCGA) (<https://cancergenome.nih.gov/>) contains patient survival and RNA sequencing data from 196 cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) human patient samples. TCGA-CESC reads per kilobase million (RPKM) data was extracted, and average RPKM for all patients combined was calculated. TCGA also contains RPKM data from 3 individual non-cancerous cervical samples that were averaged together to compare to CESC RPKM values. The Atlas of non-coding RNA in Cancer (TANRIC) (MD Anderson Cancer Center)⁴ that utilizes TCGA-CESC dataset to characterize the expression profiles of long non-coding RNAs (lncRNAs) was used to extract FAM83H-AS1 expression in each of the TCGA-CESC patient samples. Classification and Regression Trees (CART) analysis was used to statistically define high/low lncRNA expression groups. The cut-off z score value was 3 to discern between high and low expression groups. Survival curves were estimated by the Kaplan-Meier method. The log-rank test was used to estimate statistical differences between survival curves. GraphPad software was used to make the survival plots.

O. Statistics

Student's t-test was utilized to determine mean values differences between groups examined and significance was determined at $p \leq 0.05$ (*) and $p \leq 0.01$ (**). Scale bars represent Standard Deviations (SD).

XIV. Supplemental References

- 1 Liu, X. *et al.* ROCK inhibitor and feeder cells induce the conditional reprogramming of epithelial cells. *The American journal of pathology* **180**, 599-607, doi:10.1016/j.ajpath.2011.10.036 (2012).
- 2 Schowalter, R. M., Pastrana, D. V. & Buck, C. B. Glycosaminoglycans and sialylated glycans sequentially facilitate Merkel cell polyomavirus infectious entry. *PLoS pathogens* **7**, e1002161, doi:10.1371/journal.ppat.1002161 (2011).
- 3 Martinez, I. *et al.* An Exportin-1-dependent microRNA biogenesis pathway during human cell quiescence. *Proceedings of the National Academy of Sciences of the United States of America* **114**, E4961-e4970, doi:10.1073/pnas.1618732114 (2017).
- 4 Li, J. *et al.* TANRIC: An Interactive Open Platform to Explore the Function of lncRNAs in Cancer. *Cancer research* **75**, 3728-3737, doi:10.1158/0008-5472.can-15-0273 (2015).

Chapter 3

An Exportin-1-dependent microRNA biogenesis pathway during human cell quiescence

Ivan Martinez,^{a,1,2} Karen E. Hayes,^{a,1} Jamie A. Barr,^a Abby D. Harold,^a Mingyi Xie,^b Syed I. A. Bukhari,^c Shobha Vasudevan,^c Joan A. Steitz,^{d,e,f,2} and Daniel DiMaio^{d,f,g,h}

^aDepartment of Microbiology, West Virginia University Cancer Institute, West Virginia University, Morgantown, WV, 26506; ^bDepartment of Biochemistry and Molecular Biology, University of Florida Health Cancer Center, University of Florida, Gainesville, FL, 32610; ^cCancer Center, Massachusetts General Hospital, Harvard Medical School, Boston, MA, 02114; ^dDepartment of Molecular Biophysics and Biochemistry, Yale University, New Haven, CT, 06536; ^eHoward Hughes Medical Institute, Yale University, New Haven, CT, 06536; ^fYale Cancer Center, New Haven, CT, 06520; ^gDepartment of Genetics, Yale School of Medicine, New Haven, CT, 06510; ^hDepartment of Therapeutic Radiology, Yale School of Medicine, New Haven, CT, 06510

Key Words: XPO5, XPO1, (TMG)-cap, pri-miRNA, quiescence

¹Authors contributed equally to this work.

²**Correspondence:** Ivan Martinez (ivmartinez@hsc.wvu.edu) and Joan A. Steitz (joan.steitz@yale.edu)

Published in: Proceedings of the National Academy of Sciences in the United States of America
Proc Natl Acad Sci U S A. 2017 Jun 20;114(25):E4961-E4970. PMID: 28584122

I. Contribution to the Project

I assisted in numerous ways to progress the research presented in Chapter 3 to publication. I primarily conducted experiments that advanced the content in Figures 3-6. While obtaining data for these figures, I cultured cells (including human foreskin fibroblasts (HFFs) and human embryonic kidney (HEK293T)), and, on multiple occasions, I assisted with plating, inducing cellular quiescence (by serum starvation or confluency), transfecting, transducing, and pelleting cells, as needed (in experiments not described below). Additionally, I contributed to the development of several experiments stated below.

I conducted preliminary experiments in HFFs to optimize transfection conditions for future experiments. I determined optimal amounts of plasmid (GFP) and Lipofectamine for proliferating and quiescent cells (data not shown).

I conducted a preliminary experiment to determine how long microRNAs (miRNAs) could be expressed after knockdown of Exportin-1 (XPO1) followed by induction of quiescence for future experiments. To do this, I cultured and transfected HFF cells with Lipofectamine and siRNA against XPO1. Once cells were 100% confluent, I pelleted the cells at Day 0, 3, 6, 9, 12 for protein and RNA expression analysis. I conducted RNA extraction, cDNA synthesis, and qRT-PCR analysis to confirm knockdown of XPO1 and differential expression of other select genes (data not shown).

We analyzed downstream effects of individual knockdown of Exportin-5 (XPO5) (Fig. 3A, *SI Appendix*, Figs. S6A) and XPO1 (Figs. 3B-C, *SI Appendix*, Figs. S6B and S7A-E) in proliferating and quiescent HFFs using multiple siRNAs against XPO5 and XPO1, respectively. The Western Blot analysis I conducted were used in the manuscript for siRNA against XPO5 (Fig. 3A (right panel), *SI Appendix* Fig. S6A (middle panel)) and XPO1 (Fig. 3B (right panel), *SI Appendix*, Figs. S6B (middle panel) and S7D) as well as the samples I conducted RNA extraction, cDNA synthesis, and qRT-PCR analysis on to show RNA changes with individual siRNA against XPO5 (*SI Appendix*, Fig. S6A (left panel)) and individual siRNA against XPO1 (*SI Appendix*, Fig. S6B (left panel), S7A-C). Collaborators then used these samples for downstream experiments (Fig. 3C, *SI Appendix*, Figs. S6A-B (right panels), *SI Appendix*, Fig. S7E).

We exogenously expressed XPO1 in HFFs, immunoprecipitated XPO1, and discovered primary microRNAs (pri-miRNAs) that are bound to XPO1 (Fig. 3D, *SI Appendix*, Figs. S7F, S7G). To do this, we initially optimized transfection conditions (plasmid concentration and Lipofectamine volume) of an expression vector (not retroviral) (p3XFlag-XPO1). A collaborator transfected the cells. I then pelleted and

conducted RNA extraction, cDNA synthesis, and qRT-PCR analysis as well as protein extraction and Western Blot analysis to confirm XPO1 expression (data not shown). We were having difficulty exogenously expressing XPO1 with this construct, so I transformed and maxi-prepped to extract fresh plasmid. I performed a restriction digestion and ran the samples in an agarose gel to confirm the correct predicted size insert was digested (data not shown). To quickly determine if fresh plasmid solved the issues, I transfected the p3XFlag-XPO1 plasmid into HEK293T cells with Lipofectamine and performed Western Blot analysis (data not shown). A collaborator attempted to transfect the XPO1 expression constructs by nucleofection, and I performed Western Blot analysis to detect XPO1 protein knockdown (data not shown). Then, a collaborator transfected proliferating and serum-starved HFFs with p3X-Flag XPO1 and attempted multiple immunoprecipitations (IP) with an antibody against Flag. For this IP, I aided by crosslinking and conducting downstream RNA extraction, cDNA synthesis, and qRT-PCR analysis as well as protein extraction and Western Blot analysis (data not shown). Due to technical issues, a collaborator cloned XPO1 into a retroviral vector (pBABE-Flag-XPO1). I transfected HEK293T with a calcium phosphate protocol to produce a retrovirus that was transduced into HFFs to successfully exogenously expressed XPO1 in HFFs. A collaborator serum-starved cells, then I crosslinked, immunoprecipitated with an antibody against Flag, performed Western Blot analysis (*SI Appendix*, Fig. S7F (left panel)) and initiated RNA extraction used for the figure in the manuscript to identify pri-miRNAs bound to XPO1 (Fig. 3D, *SI Appendix*, Fig. S7F (right panel), Fig. S7G). I performed replicates of this experiment. It was desired for potential future experiments to stably select a pool population and single cell clones of HFFs expressing XPO1. Unfortunately, after I transfected HEK293T (pBABE-Flag-XPO1), transduced HFFs, selected with antibiotics, and cultured the "stable" pool and single cell clones, they did not maintain exogenous expression of XPO1 after multiple passages (data not shown).

We knocked down TGS1 in proliferating and quiescent HFFs using siRNA against TGS1 (Fig. 4, *SI Appendix*, Figs. S8 and S9). Prior to my assistance, my collaborators were observing only moderate knockdown of TGS1 at the RNA level with siRNA against TGS1 in HFFs. To aid in improving TGS1 knockdown, we performed optimization experiments. I varied the amount of siRNA against TGS1 (while a collaborator varied the amount of Lipofectamine concentration) transfected into proliferating and quiescent cells to determine the optimal conditions to observe TGS1 knockdown without toxicity to the cells. I conducted RNA extraction, cDNA synthesis, and qRT-PCR analysis as well as protein extraction and Western Blot analysis on all of our samples to optimize TGS1 knockdown efficiency in HFFs (data not shown). After optimization, I significantly aided in conducting multiple protein extraction and Western Blot analysis experiments as well as multiple RNA extraction, cDNA synthesis, and qRT-PCR analysis

using pellets provided by a collaborator who performed the upstream portion of experiments (Fig. 4, *SI Appendix* Figs. S8 and S9).

After we obtained results presented in Figure 4, we pursued the potential of Snurportin-1 in regulating localization of the TMG-capped pri-miRNAs induced during quiescence (data not shown). (Snurportin-1 is involved with recognizing a (m^{2,2,7}G, TMG)-cap on certain small nucleolar RNAs (snRNAs) to import them to the nucleus from the cytoplasm(1)). I performed preliminary experiments to explore the regulation of Snurportin-1 in quiescent versus proliferating cells. To do this, I conducted Western Blot analysis of HFFs cells that were quiescent for 0, 1, 2, 3 days to analyze for protein expression changes in Snurportin-1 (data not shown). Additionally, I conducted Western Blot analysis on HFF cell pellets with siRNA against Snurportin-1 provided by a collaborator so I could detect knockdown of Snurportin-1 protein (data not shown).

We knocked down Drosha and XPO1 individually as well as together in proliferating and quiescent HFFs for downstream experiments (Fig. 5, *SI Appendix*, Figs. S10 and S12). I conducted Western Blot analysis to confirm Drosha and/or XPO1 knockdown (*SI Appendix*, Fig. S12 (top panel)) prior to a collaborator performing downstream experiments with these samples (Fig. 4, *SI Appendix*, Figs. S12 (bottom panel), S10).

After we identified an increased expression of a short isoform of Drosha in quiescent cells (Fig. 6A), we used a pool of three siRNAs to target Drosha mRNA in proliferating and quiescent cells and determined that these siRNAs were knocking down both isoforms of Drosha (*SI Appendix*, Fig. S13A). To do this experiment, a collaborator performed upstream knockdown experiments, and I processed the provided pellets through protein extraction and Western Blot analysis. The data I obtained was used in the manuscript (*SI Appendix*, Fig. S13A).

We performed cellular fractionation of quiescent and proliferating HFFs to determine localization of Drosha (Fig. 6B). While a collaborator was optimizing cellular fractionation conditions for proliferating and serum-starved HFFs, I processed multiple rounds of pellets for protein extraction and Western Blot analysis. Additionally, I obtained the data that was used in the manuscript (Fig. 6B).

Moreover, I proofread the manuscript prior to the initial submission and re-submission after revisions as well as proofread the comments to the reviewers upon re-submission.

In summary, my experimental and writing contribution to this specific manuscript were substantial. The participation in this project helped me learn several molecular techniques as well as the process of submission and publication of its findings.

1. **Huber J, Cronshagen U, Kadokura M, Marshallsay C, Wada T, Sekine M, Luhrmann R.** 1998. Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *Embo j* **17**:4114-4126.

II. Abstract

The reversible state of proliferative arrest known as cellular quiescence plays an important role in tissue homeostasis and stem cell biology. By analyzing the expression of microRNAs (miRNAs) and miRNA processing factors during quiescence in primary human fibroblasts, we identified a group of miRNAs that are induced during quiescence despite markedly reduced expression of Exportin-5, a protein required for canonical miRNA biogenesis. The biogenesis of these quiescence-induced miRNAs is independent of Exportin-5 and depends instead on Exportin-1. Moreover, these quiescence-induced primary (pri)-miRNAs are modified with a 2,2,7-trimethylguanosine (TMG)-cap, which is known to bind Exportin-1, and knock-down of Exportin-1 or trimethylguanosine synthase 1, responsible for (TMG)-capping, inhibits their biogenesis. Surprisingly, in quiescent cells Exportin-1-dependent pri-miR-34a is present in the cytoplasm together with a small isoform of Drosha, implying the existence of a different miRNA processing pathway in these cells. Our findings suggest that during quiescence, the canonical miRNA biogenesis pathway is down-regulated, and specific miRNAs are generated by an alternative pathway to regulate genes involved in cellular growth arrest.

III. Significance

Quiescence is a growth-arrested cellular state where genes involved in this process are finely regulated by several factors including miRNAs. During miRNA biogenesis, Exportin-5 transports miRNA precursors from the nucleus to the cytoplasm. In this study, we demonstrated the existence of an alternative miRNA biogenesis pathway in quiescent primary human cells. This pathway involves the repression of Exportin-5 expression by autophagy and miRNAs, and the (TMG)-cap modification of specific pri-miRNAs, which signals their export to the cytoplasm by Exportin-1. We further showed that these pri-miRNAs are rapidly processed in the cytoplasm by a small isoform of Drosha. Collectively, these results reveal a new mechanism of miRNA biogenesis that will expand our understanding of miRNA regulation in normal or disease-related cells.

IV. Introduction

Most metazoan cells enter a reversible cell cycle arrest known as cellular quiescence when they are exposed to anti-mitogenic signals or an environment unfavorable for proliferation (2, 3). In mammalian cells, quiescence (also known as G₀ arrest) is characterized by reduced DNA replication, altered metabolism, increased autophagy, and increased expression of cyclin-dependent kinase inhibitors such as

p27^{Kip1} (4, 5). *In vitro*, quiescence can be induced in primary cells by serum starvation, contact inhibition, and loss of adhesion to a substrate (6). Quiescence is involved in important cellular processes including the balance between differentiation and self-renewal in different types of stem cells, and dysregulation of quiescence could favor carcinogenesis. However, the molecular mechanisms that regulate quiescence are poorly understood (7).

MicroRNAs (miRNAs) are small non-coding RNAs ~22-nucleotides (nt) long that regulate the expression of protein-coding genes by base-pairing with the 3' untranslated region (3'UTR) of messenger RNAs (mRNAs), repressing translation and/or inducing degradation of the target mRNA (8, 9). In canonical miRNA biogenesis in mammalian cells, miRNAs are transcribed by RNA polymerase II to produce 7-methylguanosine-capped (m⁷G) primary miRNAs (pri-miRNAs) containing one or more bulged hairpin structures that are recognized by the nuclear microprocessor, which consists of the ribonuclease (RNase) III enzyme Drosha and the double-stranded RNA (dsRNA) binding protein DiGeorge syndrome critical region gene 8 (DGCR8) (10-13). Specific cleavage of the pri-miRNA by the nuclear microprocessor generates an ~70-nt stem-loop structure known as the precursor miRNA (pre-miRNA), which is recognized and transported to the cytoplasm by Exportin-5 (XPO5) (4, 11, 12, 14). Subsequently, pre-miRNA is cleaved by the cytoplasmic RNase III enzyme Dicer (15-18), generating a double-stranded mature miRNA. Finally, the guide miRNA strand is loaded into the RNA-induced silencing complex (RISC) containing one of four Argonaute proteins and GW182 protein at its core (19-21).

Recent studies have identified alternative pathways of microRNA biogenesis in different cell and animal models: pre-miRNA/introns miRNAs (miRtrons), which are Drosha and DGCR8 independent (22); Drosha and DGCR8-independent viral miRNAs that use the Integrator complex to produce viral pre-miRNAs (23); miRNAs independent of DGCR8, Exportin-5, Dicer and Argonaute 2 (AGO2) known as Simtrons (24), which are exported from the nucleus by an unknown mechanism; Dicer-independent miRNAs, such as miR-451, which are generated by the catalytic activity of AGO2 (25, 26); and cytoplasmic re-localization of Drosha by RNA viruses, with no changes in endogenous miRNA expression but robust synthesis of viral miRNAs (27). Recently, we reported a novel class of (m⁷G)-capped pre-miRNAs in human cells whose 5' end corresponds to the transcription initiation site (28); these pre-miRNAs are exported from the nucleus by Exportin-1 rather than Exportin-5.

Exportin-1 recognizes certain RNAs including viral mRNAs, cellular mRNAs, and small nuclear RNAs (snRNAs) and exports them from the nucleus to the cytoplasm (29). For example, precursors of snRNAs involved in mRNA processing such as U1, U2, U4, and U5 have a (m⁷G)-cap, which is recognized

by cap-binding complex (CBC) and the phosphorylated adaptor for RNA export (PHAX) in the nucleus to enable their export to the cytoplasm (30) by Exportin-1. These snRNAs are then recognized by Sm proteins in the cytoplasm and trimethylguanosine synthase 1 (TGS1) is recruited to hypermethylate the (m⁷G)-cap into a (m^{2,2,7}G, TMG)-cap. This modification is recognized by Snurportin-1 in association with Importin-β and other factors to import the snRNAs back into the nucleus (31, 32). Interestingly, Exportin-1 also has high affinity for the (TMG)-capped small nucleolar RNA (snoRNA) U3 in the nucleus and transports it from Cajal bodies to the nucleoli (33). Another study showed that TGS1 enhances Rev-dependent HIV-1 RNA expression by (TMG)-capping viral mRNAs in the nucleus, thereby increasing recognition by Exportin-1 for transport to the cytoplasm (34).

miRNAs play an important role in regulating cellular quiescence. For example, it has been shown that over-expression of miR-221 and miR-222, both of which target the 3'UTRs of p27 and p57 mRNA, induce S-phase entry in quiescent cells (35). Another study reported up-regulation of let-7 and miR-125 and down-regulation of the miR-29 family in quiescent fibroblasts (36). Furthermore, over-expression of miR-29 increased cell cycle re-entry in quiescent cells, whereas over-expression of let-7 and miR-125 delayed cell cycle re-entry (36). Modulation of miR-29 can also influence senescence, an irreversible growth arrested state, by affecting B-myb expression (37). We also showed that during quiescence some miRNAs up-regulate translation of their target mRNAs (38). Finally, Exportin-5 protein expression is reduced during quiescence and promptly induced during cell cycle entry, causing a global increase in miRNA expression (39).

Here, we report the existence of an alternative miRNA biogenesis pathway in primary human foreskin fibroblasts (HFFs) during quiescence. We confirmed that the level of Exportin-5 is substantially reduced during quiescence and that certain miRNAs are induced during quiescence despite this paucity of Exportin-5. We also demonstrate that the biogenesis of pri-miRNAs corresponding to these miRNAs is dependent not on Exportin-5, but rather Exportin-1. Furthermore, the Exportin-1-dependent pri-miRNAs have a trimethylguanosine (m^{2,2,7}G, TMG)-cap and are found in the cytoplasm during quiescence, as is a smaller isoform of Drosha. Together, our findings suggest that in quiescent cells, an alternative miRNA biogenesis pathway selectively processes and transports a specific set of miRNAs, which could be essential for reversible G₀ arrest.

V. Results

A. Differential expression of specific pri-miRNAs and their corresponding mature miRNAs during quiescence

To examine global miRNA expression during quiescence, miRNA microarrays that interrogated 1105 mature human miRNAs were used to compare proliferating HFFs and quiescent HFFs after 12, 48, and 72 h of serum starvation. We identified 88 up-regulated and 94 down-regulated miRNAs that changed 1.5-fold or more for at least two time-points of serum starvation (Fig. 1A and *SI Appendix*, Tables S1, S2). As assessed by qRT-PCR, mature miR-26a-2, miR-34a, miR-126, miR-199b, miR-638, miR-1228*, and miR-3188 were induced two- to four-fold by serum starvation, and miR-17, miR-18a, miR-29b, miR-155, and miR-423-3p were repressed two- to three-fold, confirming the microarray data (Fig. 1B). There was a similar pattern of miRNA expression (with the exception of miR-26a-2 and miR-126) when quiescence was induced by confluency (*SI Appendix*, Fig. S1A). Northern blot analysis confirmed induction of representative miRNAs in quiescence induced by serum starvation or confluency (*SI Appendix*, Fig. S2). By using specific primers to amplify the pri-miRNAs of these differentially expressed miRNAs (*SI Appendix*, Table S3), we found that the mature miRNAs induced during quiescence are also induced at the pri-miRNA level (Fig. 1C and *SI Appendix*, Fig. S1B). In contrast, rather than being repressed, the levels of the pri-miRNAs corresponding to the repressed mature miRNAs miR-17, miR-18a, miR-29b, miR-155, and miR-423-3p were induced during quiescence (Fig. 1C and *SI Appendix*, Fig. S1B), suggesting that repression of these miRNAs is due to post-transcriptional regulation, presumably during their processing through the microRNA biogenesis pathway.

B. Exportin-5 expression is down-regulated by autophagy and miR-34a during quiescence.

To explore the mechanism of miRNA synthesis in quiescent cells, the abundance of proteins involved in miRNA biogenesis was examined in quiescent HFF. Confirming the previous findings of Iwasaki and collaborators (39), Exportin-5 protein levels were dramatically reduced (~30-fold) during quiescence induced by serum starvation (Fig. 2A, and *SI Appendix*, Fig. S3A) or contact inhibition (*SI Appendix*, Fig. S3B), as assessed by immunoblotting and immunofluorescence. Exportin-5 protein was also reduced in serum-starved HeLa (human cervical carcinoma) and C127 (murine mammary tumor) cell lines (*SI Appendix*, Fig. S3C). The level of Exportin-5 mRNA did not change during quiescence (*SI Appendix*, Fig. S3 D and E), indicating that Exportin-5 is regulated at the post-transcriptional level. The level of Exportin-1, a protein involved in protein and RNA export (29), did not change during quiescence (Fig. 2A, and *SI Appendix*, Fig. S3B). The quiescence marker p27 was induced, as expected (40).

To determine the basis for the reduction in Exportin-5 expression during quiescence, we blocked two major pathways of protein degradation, the proteasome and the autophagy pathways. Treatment of quiescent cells with the proteasome inhibitor, MG132, did not restore Exportin-5 expression, whereas the protein MDM2, which is known to be degraded by the proteasome, is maintained upon MG132 treatment, as expected (Fig. 2B, and *SI Appendix*, Fig. S4A). In fact, the level of Exportin-5 was further reduced in response to MG132 treatment, suggesting that a down-regulation pathway is induced by this drug (Fig. 2B). Treatment with 3-methyladenine (3-MA), an autophagy inhibitor, reduced the level of Exportin-5 in proliferating cells (Fig. 2C). Strikingly, treatment of quiescent HFFs with 3-MA restored the levels of Exportin-5 and of p62, a protein specifically degraded by autophagy (Fig. 2C and *SI Appendix*, Fig. S4B). These data suggest that the autophagy pathway is responsible for accelerated degradation of Exportin-5 during quiescence.

We also tested whether Exportin-5 levels are modulated by miRNAs. The 3'UTR of the Exportin-5 mRNA contains a conserved predicted binding site for miR-34a (Fig. 2D and *SI Appendix*, Fig. S5), which as shown above is induced during quiescence (Fig. 1B and *SI Appendix*, Fig. S1A). To determine if miR-34a affects expression of Exportin-5, we transfected synthetic miR-34a into proliferating HFF cells and examined the expression of Exportin-5 after 48 h. As shown in Figure 2E, transfected miR-34a reduced the amount of endogenous Exportin-5 protein but not Exportin-5 mRNA, suggesting translational repression of Exportin-5 by miR-34a. In contrast, there was no change in expression of Exportin-1, which does not contain a miR-34a binding site in its 3'UTR. To determine if there is a direct functional interaction between miR-34a and Exportin-5 mRNA, we cloned the complete 3'UTR of the Exportin-5 gene (1,537 nt) into a luciferase reporter vector and transfected quiescent HFF cells with the empty reporter vector, the reporter vector containing the Exportin-5 3'UTR, or a reporter vector containing the Exportin-5 3'UTR with three point mutations in the miR-34a seed sequence binding site (Fig. 2D). As shown in Figure 2F, luciferase activity from the vector containing the wild-type Exportin-5 3'UTR was reduced about two-fold during quiescence compared to the vector lacking the 3'UTR, and there was partial recovery of luciferase activity in response to the mutant 3'UTR. These results suggest that regulation of Exportin-5 expression in quiescent cells is mediated in part by miR-34a. We also co-transfected proliferating HFFs with the Exportin-5 3'UTR luciferase reporter vectors and synthetic miR-34a or scrambled control miRNA. miR-34a but not the control miRNA caused a reduction in luciferase activity from the wild-type Exportin-5 3'UTR reporter but not from the miR-34a binding site mutant or empty vector (Fig. 2G). Together these results demonstrate a functional interaction between Exportin-5 3'UTR and miR-34a. We conclude that Exportin-5 expression is regulated at two different levels during quiescence, at the protein level by autophagy and at the mRNA level by miR-34a.

C. Expression of quiescence-induced miRNAs is not dependent on Exportin-5.

The increase in the levels of specific mature miRNAs in quiescent cells expressing little Exportin-5 raised the possibility that these quiescence-induced miRNAs are generated by an Exportin-5-independent pathway. To investigate this possibility, we transiently transfected proliferating HFFs with an siRNA against Exportin-5 and 48 h later measured miRNA levels. Exportin-5 knock-down was confirmed by Western blot and qRT-PCR (Fig. 3A). As expected, Exportin-5 knock-down in proliferating cells lowered levels of mature miR-17, miR-18a, miR-29b, miR-155, and miR-423-3p, which are repressed during quiescence (Fig. 3A, dark gray bars). In contrast, levels of their corresponding primary miRNAs did not decline (Fig. 3A, light gray bars), consistent with Exportin-5 being required for maturation of these miRNAs, as it is for all canonically processed miRNAs. Strikingly, however, Exportin-5 knock-down did not repress the levels of several mature miRNAs and pri-miRNAs that are induced during quiescence, including miR-26a, miR-34a, miR-126, miR-199b, miR-638, and miR-3188 (Fig. 3A). Similar results were obtained with a second Exportin-5 siRNA in an independent experiment (*SI Appendix*, Fig. S6A). Although, we cannot exclude other possibilities such as a higher affinity for Exportin-5 by some quiescence-induced miRNAs (which can thus tolerate lower Exportin-5 levels), or differences in the stability of the mature miRNA (41), our findings suggest that the biogenesis of these miRNAs can proceed via an Exportin-5-independent mechanism, even in proliferating cells.

D. Exportin-1 is required for processing of quiescence-induced miRNAs.

The results in the preceding section imply that an export factor other than Exportin-5 is required for the synthesis of mature miRNAs during quiescence. Previous studies showed that Exportin-1, an essential transport factor for certain proteins and RNAs, supports pri-miRNA processing in *C. elegans* and *Drosophila*, as well as shuttling of mature miRNA from the cytoplasm to the nucleus (29, 42, 43). Recently, we showed that Exportin-1 is involved in the nuclear export of a group of mammalian (m⁷G)-capped pre-miRNAs in proliferating cells (28).

To test whether Exportin-1 is involved in the biogenesis of the miRNAs induced in quiescent cells, we transiently transfected proliferating HFFs with an siRNA targeting Exportin-1 and confirmed siRNA-mediated reduction of Exportin-1 but not Exportin-5 expression (Fig. 3B). Strikingly, Exportin-1 knock-down in proliferating cells inhibited the expression of several of the mature miRNAs that are induced during quiescence, namely miR-26a, miR-34a, miR-126, miR-199b, miR-638, miR-1228*, and miR-3188 (Fig.

3B, dark gray bars). Notably, the levels of the pri-miRNAs corresponding to these miRNAs did not decline (Fig. 3B, light gray bars); in fact, pri-miR-34a increased five- to nine-fold in the absence of Exportin-1. The levels of Exportin-5-dependent miRNAs miR-17, miR-18a, miR-29b, miR-155, and miR-423-3p were also reduced in proliferating cells by Exportin-1 knock-down (Fig. 3B, dark gray bars), but the corresponding primary miRNAs of these mature miRNAs were also decreased, suggesting a reduction of transcription or changes in pri-miRNA turnover in the absence of Exportin-1. Mature miR-155 and the Simtron miR-1228* but not their corresponding pri-miRNAs were reduced by both Exportin-5 and Exportin-1 knock-down (Fig. 3 A and B, dark gray bars), suggesting that both exportins are involved in the biogenesis of these miRNAs. Similar results were obtained with a second Exportin-1 siRNA experiment (*SI Appendix*, Fig. S6B).

To test the contribution of Exportin-1 to miRNA levels during quiescence, we knocked down Exportin-1 in serum-starved cells and measured the levels of mature miRNAs. Exportin-1 knock-down in quiescent cells did not affect the level of Exportin-5 or p27 mRNA or other proteins involved in miRNA biogenesis such as DGCR8, Drosha, Exportin-5, Dicer, and AGO2 (*SI Appendix*, Fig. S7A-D). Nevertheless, Exportin-1 knock-down prevented or attenuated the increase of the quiescence-induced miRNAs: miR-34a, miR-638, miR-3188, and miR-1288* (Fig. 3C). The Exportin-1-dependence of miR-34a expression in quiescent cells was confirmed by Northern blot (Fig. 3C). Because the expression of the other quiescence-induced miRNAs was difficult to detect by Northern blotting, we focused on miR-34a in subsequent experiments. Interestingly, Exportin-1 knock-down did not change the percentage of HFFs entering quiescence after 24 h of serum starvation but caused a two-fold decrease in the percentage of cells in S phase 24 h after the addition of serum, suggesting that Exportin-1-mediated miRNA biogenesis may be required for normal exit from quiescence (*SI Appendix*, Fig. S7E).

We used protein-RNA crosslinking followed by Exportin-1 immunoprecipitation to determine if there is a physical interaction between Exportin-1 and pri-miR-34. For this experiment we stably expressed FLAG-tagged Exportin-1 and immunoprecipitated cell extracts with an antibody recognizing the FLAG epitope (*SI Appendix*, Fig. S7F). The control snoRNA U3, which is normally transported by Exportin-1 (33), was co-immunoprecipitated as expected (Fig. 3D and *SI Appendix*, Fig. S7G). Strikingly, the Exportin-1-dependent pri-miR-34a but not the Exportin-5-dependent pri-miR-423-3 was co-immunoprecipitated from quiescent HFFs by anti-Flag-Exportin-1 (Fig. 3D, and *SI Appendix*, Fig. S7G). Together, these results suggest that Exportin-1 is required for the export and/or processing of miRNAs induced during quiescence and that this activity involves a physical association between these pri-miRNAs and Exportin-1.

E. Exportin-1-dependent pri-miRNAs have a (TMG)-cap during quiescence.

Previous reports showed that Exportin-1 can bind to (TMG)-capped RNAs such as snoRNA U3 and viral Rev-dependent HIV-1 RNAs (33, 34). Therefore, we investigated whether pri-miRNAs of the Exportin-1-dependent miRNAs contain a (TMG)-cap. We used an antibody against (TMG)-caps (which does not cross-react with the (m⁷G)-cap (44)) and non-immune rabbit serum to perform immunoprecipitation from extracts of proliferating or quiescent HFFs. Specific immunoprecipitated pri-miRNAs were detected by RT-PCR amplification followed by gel electrophoresis. As shown in Figure 4A, top panel, the control (TMG)-capped U3 snoRNA was amplified following anti-(TMG)-cap pull-down from both proliferating and quiescent HFF, as expected. Strikingly, pri-miR-34a and pri-miR-3188, two quiescence-induced miRNAs, were also detected in the anti-(TMG)-cap immunoprecipitate from quiescent and to a lower extent from proliferating HFFs. qRT-PCR data from a replicate experiment confirmed that pri-miR-34a and pri-miR-3188 were immunoprecipitated by the anti-(TMG)-cap antibody during proliferation and quiescence (*SI Appendix*, Fig. S8 and *SI Appendix*, Table S4). In contrast, pri-miR-29b and pri-miR-423, which are repressed during quiescence, were not detectable in the anti-(TMG)-cap pull-downs. These results indicate that some quiescence-induced pri-miRNAs are specifically modified with a (TMG)-cap.

TGS1, the enzyme that catalyzes 5'-cap hypermethylation, has two active isoforms, a short isoform in the nucleus and a full-length isoform in the cytoplasm (45). We found that the short form of TGS1 is induced during quiescence, suggesting it may be responsible for cap hypermethylation of Exportin-1-dependent pri-miRNAs (*SI Appendix*, Fig. S9A). Indeed, the ability of the anti-TMG antibody to immunoprecipitate pri-miR-34a and pri-miR-3188 (as well as the control U3 RNA) was dramatically diminished by siRNA-mediated knock-down of TGS1 in proliferating and quiescent HFFs (Fig. 4A, compare siRNA CTRL anti-TMG IP lanes against siRNA TGS1 anti-TMG IP lanes, and *SI Appendix*, Fig. S8 and S9 B and C), arguing that this enzyme is required for cap hypermethylation of these miRNAs.

To further test the model that TMG capping is required for miRNA biogenesis in quiescent cells, we examined levels of mature miRNAs after knock-down of TGS1 in quiescent HFFs. As shown in Figure 4B, TGS1 knock-down inhibited the induction of the quiescence-induced miR-34a and miR-3188 but did not affect the abundance of the quiescence-repressed miR-423-3p and miR-29b. The knock-down of TGS1 did not affect the expression of DGCR8, Drosha, or Exportin-5, and actually increased the expression of Dicer and AGO2 (*SI Appendix*, Fig. S9D). To investigate whether the (TMG)-cap modification in quiescence-induced pri-miRNAs is important for the interaction with Exportin-1, we used protein-RNA

crosslinking followed by immunoprecipitation of FLAG-Exportin-1 in quiescent HFFs that had been treated with control scrambled siRNA or siRNA against TGS1. snoRNA U3, which contains a (TMG)-cap, showed a reduced level of co-immunoprecipitation from cells treated with siRNA against TGS1, as expected (Fig. 4C). Notably, Exportin-1-dependent pri-miR-34a was also co-immunoprecipitated to a lesser extent in cells after TGS1 knock-down (Fig. 4C), whereas the Exportin-5-dependent pri-miR-423 was not co-immunoprecipitated in the presence or absence of TGS-1, reaffirming the hypothesis that this pri-miRNA does not contain a modified cap.

Taken together, these findings indicate that the pri-miRNAs of quiescence-induced miRNAs are modified with a specific hypermethylation in their 5'-cap by TGS1, a modification important for Exportin-1 recognition (33).

F. Pri-miR-34a is present in the cytoplasm of quiescent cells together with a small isoform of Drosha.

To examine whether quiescence-induced pri-miRNAs are exported to the cytoplasm as a consequence of (TMG)-capping and interaction with Exportin-1, we prepared nuclear and cytoplasmic RNA from proliferating and quiescent HFFs pretreated with control siRNA or a pool of siRNAs against Drosha (in order to minimize possible rapid processing of these pri-miRNAs (*SI Appendix*, Fig. S10)). As expected, Exportin-1-dependent pri-miR-34a and Exportin-5-dependent pri-miR-29b and pri-miR-423 were detected in the nuclear fraction of both proliferating and quiescent HFFs (Fig. 5A). These nuclear pri-miRNAs accumulated in response to Drosha knock-down, supporting previous studies that showed their Drosha dependency (46, 47). Strikingly, we also detected pri-miR-34a in the cytoplasmic fraction of proliferating and quiescent HFFs following Drosha knock-down, whereas pri-miR-29b and pri-miR-423 were not detectable in this cellular fraction (Fig. 5 A and B). These data suggest that pri-miR-34a and possibly other quiescence-induced pri-miRNAs are transported into the cytoplasm where they are processed by an uncharacterized cytoplasmic form of Drosha. We also observed a reduction of all the pri-miRNAs in the nuclear fraction of the quiescent cells (Fig. 5A and *SI Appendix*, Fig. S11).

To determine if the presence of cytoplasmic pri-miR-34a required Exportin-1, we simultaneously knocked down Exportin-1 as well as Drosha in quiescent HFFs and determined the level of cytoplasmic pri-miRNAs compared to cells treated with control siRNA or Drosha siRNA alone (*SI Appendix*, Fig. S12). As shown in Figure 5C and D, knock-down of Exportin-1 expression in quiescent HFFs greatly attenuated the increased amount of cytoplasmic pri-miR-34a in the Drosha knockdown cells and caused the

accumulation of this pri-miRNA in the nuclear fraction, suggesting that Exportin-1 is involved in the transport of pri-miR34a from the nucleus to the cytoplasm.

To further explore the processing of pri-miRNAs in quiescent HFFs, we analyzed Drosha protein during quiescence. Western blotting showed the presence of full-length Drosha (~160 kDa) in proliferating and quiescent cells (Fig. 6A). In addition, quiescent cells displayed induction of a smaller Drosha band (~130 kDa) (Fig. 6A). A pool of three siRNAs that target Drosha mRNA in different regions caused a reduction in both bands, suggesting that the smaller band is a distinct splicing isoform or cleavage product of Drosha (*SI Appendix*, Fig. S13 A and B). An earlier publication also detected a smaller form of Drosha (~145 kDa) in HEK-293 cells, which was observed in a different protein complex than full-length Drosha, suggesting that it associates with different binding partners (12).

To investigate the localization of these two different forms of Drosha, we examined nuclear and cytoplasmic fractions from proliferating and quiescent HFFs. The short form of Drosha was found specifically in the cytoplasmic fraction of quiescent HFFs (Fig. 6B). In addition, immunofluorescent staining of Drosha in proliferating and quiescent HFFs revealed a cytoplasmic signal during quiescence (Fig. 6 C and D), which was reduced by siRNA against Drosha (*SI Appendix*, Fig. S13E). Because DGCR8 is essential for Drosha activity, we analyzed the cytoplasmic fraction of proliferating and quiescent HFFs for DGCR8. We were unable to detect DGCR8 in the cytoplasmic fractions by direct Western blotting (Fig. 6B) or by blotting after immunoprecipitation (*SI Appendix*, Fig. S14A), suggesting that Drosha may interact with a different dsRNA binding protein in the cytoplasm, as was previously suggested (27). It is noteworthy that two recent publications demonstrated that alternative splicing of Drosha could affect its subcellular localization, consistent with our finding that a specific Drosha isoform exists in the cytoplasm of quiescent cells (48, 49). Furthermore, siRNA against DGCR8 in quiescent HFFs did not affect the biogenesis of the quiescence-induced miR-34a but altered the processing of the quiescence-repressed miR-423-3p (*SI Appendix*, Fig. S14B), supporting the existence of functional cytoplasmic Drosha without DGCR8.

To characterize the short form of Drosha in quiescent cells, we amplified the Drosha open reading frame mRNA by tiling RT-PCR (*SI Appendix*, Fig. S15). Cloning and sequencing identified two exon-skipping events in the same transcript, occurring in the 5' half Drosha mRNA, during quiescence. These events disrupt the proline-rich (P-rich) region as well as the nuclear localization signal in the arginine and serine-rich (RS-rich) portion of Drosha (Fig. 6E), but leave intact the previously described minimal active fragment of the protein (18). To indirectly measure the activity of the short isoform of Drosha, we measured the levels of pri-miR-34a and pri-miR-423 after treating cells with siRNAs against both Drosha isoforms

(siRNA Drosha both) or an siRNA that specifically targets the exon 6 region that is present only in the full-length Drosha but not in the short isoform (siRNA Drosha full-length) (*SI Appendix*, Fig. S13 C and D). siRNA against both forms of Drosha caused a marked increase in the level of pri-miR-34a and pri-miR423 in quiescent cells, confirming that Drosha contributes to the processing of these miRNAs (Fig. 6F). Notably, siRNA specifically targeting full-length Drosha was much less effective in inducing pri-miR-34a or miR-34a in quiescent cells than was siRNA in reducing both Drosha isoforms, implying that the short cytoplasmic form of Drosha was competent to process pri-miR-34a to mature miR-34a. In contrast, the short isoform of Drosha does not appear to be involved in processing pri-miR-423 because specific knock-down of the full-length form was sufficient to cause pri-miR-423 accumulation. Similarly, siRNA against full-length Drosha allows partial induction of mature miR-34a but not miR-423 in quiescent cells (Fig. 6G), supporting the hypothesis that the short isoform of Drosha is able to process pri-miR-34a but not pri-miR-423 during quiescence.

VI. Discussion

When cells reversibly exit the cell cycle and enter quiescence, they remain in a non-proliferative state until they receive appropriate mitogenic signals, but must avoid irreversible states such as apoptosis, differentiation, and senescence. Because miRNAs regulate many cellular pathways, changes in miRNA biogenesis are likely to play an important role in controlling these important cellular processes. This study revealed a number of unique features of miRNA biogenesis in proliferating and quiescent cells.

A. Exportin-5 levels are regulated during cellular quiescence by autophagy and miR-34a.

We confirmed the findings of Iwasaki and collaborators (39) that the amount of Exportin-5 protein decreases during quiescence (Fig. 2A and *SI Appendix*, Fig. S3). We showed that at least two post-transcriptional mechanisms contribute to the reduction of Exportin-5 expression during quiescence. Because a specific inhibitor of autophagy prevented the loss of Exportin-5, a major regulator of Exportin-5 levels during quiescence appears to be protein degradation by autophagy. In fact, autophagy is induced during quiescence to protect cells from apoptosis mediated by the proteasome (50), and serum starvation or treatment with an mTOR inhibitor activates autophagy, resulting in the degradation of Dicer and AGO2 (51). Together with our findings, these results indicate that several proteins involved in miRNA biogenesis are regulated by autophagy during quiescence to avoid induction of apoptosis. In addition, our studies suggest that miR-34a binds to an evolutionarily conserved binding site in the Exportin-5 mRNA 3'UTR and represses Exportin-5 expression (Fig. 2D-G). miR-34a-mediated repression of Exportin-5 appears to

act in conjunction with autophagy-mediated degradation to ensure low levels of Exportin-5 throughout quiescence. Consistent with this idea, miR-34a is induced relatively late after the onset of quiescence, when Exportin-5 levels are already low (Fig. 2A and *SI Appendix*, Fig. S3). Taken together, our data suggest that Exportin-1 contributes to Exportin-5 down-regulation during quiescence by supporting the synthesis of miR-34a, which binds to Exportin-5 mRNA and reinforces the block of Exportin-5 expression. An interesting topic for future studies will be to determine if the regulation of Exportin-5 is related exclusively to the induction of quiescence or rather to the absence of specific growth factors such as IGF-1, which have been shown to be important in the regulation of autophagy in HFFs (52).

B. Exportin-1-dependent microRNA biogenesis occurs during cellular quiescence.

Our major finding is the discovery of an alternative miRNA biogenesis pathway involving Exportin-1 (also known as CRM1) during quiescence. Exportin-1, the major factor responsible for the export of proteins from the nucleus, is also involved in the export of certain RNAs such as ribosomal RNAs, small nuclear RNAs, and mRNAs encoding some proteins involved in cell cycle regulation (29, 53). Previous studies in *C. elegans* and *Drosophila* showed that Exportin-1 is involved in pri-miRNA processing and biogenesis (42). Our findings demonstrate that Exportin-1 is also involved in the biogenesis of specific miRNAs in proliferating mammalian cells and that this activity is enhanced during quiescence. First, there is no change in Exportin-1 levels during quiescence (Fig. 2A and *SI Appendix*, Fig. S3B). Second, Exportin-1 knock-down reduced the levels of specific quiescence-induced mature miRNAs in proliferating and quiescent HFFs (Fig. 3 B and C), although the effect was not as dramatic as the reduction of quiescence-repressed miRNAs by Exportin-5 knock-down (Fig. 3A). Third, the Exportin-1-dependent pri-miR-34a but not the Exportin-5-dependent pri-miR-423 is found in a stable complex with Exportin-1 (Fig. 3D). Our data further demonstrate that some miRNAs, including miR-26a, miR-34a, miR-126, miR-199, miR-638, and miR-3188, can use an Exportin-5-independent biogenesis pathway in proliferating cells. We also identified miRNAs subject to complex regulation during quiescence. For example, siRNA-mediated knock-down of either Exportin-5 or Exportin-1 in proliferating HFFs reduced the expression of mature miR-1228* but did not affect levels of pri-miR-1228 (Fig. 3 A and B, and *SI Appendix*, Fig. S6 A and B). Thus, both Exportin-5 and Exportin-1 are involved in the biogenesis of miR-1228*, which is considered a Simtron, a class of miRNAs that are produced in the absence of splicing and are independent of DGCR8, Dicer, Exportin-5, and Argonaute (24). Kim and collaborators (48) demonstrated that Exportin-5 knockout in the colorectal cancer cell line HCT116 only modestly affects the processing of several miRNAs, consistent with the existence of Exportin-5-independent miRNA maturation in these cells.

C. Quiescence-induced primary miRNAs contain a (TMG)-cap and bind Exportin-1.

We discovered that the primary miRNAs of Exportin-1-dependent quiescence-induced miRNAs contain a (TMG)-cap, which may be a signal for recognition by Exportin-1, whereas this cap structure is absent from the Exportin-5-dependent pri-miRNAs (Fig. 4A, *SI Appendix*, Fig. S8, and *SI Appendix*, Table S4). We also found that a short nuclear isoform of TGS1 is up-regulated during quiescence (*SI Appendix*, Fig. S9A). Importantly, knocking down expression of TGS1 in HFFs reduced (TMG)-capping of quiescence-induced pri-miRNAs, diminished induction of the corresponding mature miRNAs during quiescence, and reduced the interaction of pri-miR-34a with Exportin-1 (Fig. 4). Because most RNAs need to be exported to the cytoplasm to be modified with a (TMG)-cap, we have not ruled out the possibility that some pri-miRNAs could be exported by Exportin-1 with a (m⁷G)-cap and then undergo cap hypermethylation in the cytoplasm. We recently described microprocessor-independent human and mouse pre-miRNAs that are specifically (m⁷G)-capped and transported by the PHAX-Exportin-1 pathway and cleaved by Dicer prior to loading into the RISC complex (28). The miRNAs described here are processed differently because TMG caps are added at the pri-miRNA level, indicating that human cells contain at least two alternative Exportin-1-dependent pathways of miRNA biogenesis.

D. Discovery of primary miRNAs and a smaller Drosha isoform in the cytoplasm of quiescent cells.

Another important discovery is finding pri-miR-34a in the cytoplasm. We detected cytoplasmic pri-miR-34a in proliferating and quiescent HFF only after knocking down Drosha, suggesting that this pri-miRNA is rapidly processed once it leaves the nucleus (Fig. 5 A and B). We were unable to consistently detect other quiescence-induced pri-miRNAs in the cytoplasm, presumably because they are less abundant than pri-miR-34a. Drosha-mediated rapid cleavage of these cytoplasmic (TMG)-capped pri-miRNAs may prevent their recognition by Snuportin-1, which could potentially return the pri-miRNAs to the nucleus (31, 32). Moreover, we observed a reduction in the amount of pri-miR-34a in the cytoplasmic fraction of cells knocked down for both Exportin-1 and Drosha (Fig. 5C). This observation, together with the discovery of an interaction between pri-miR-34a and Exportin-1 (Fig. 3D), suggests that Exportin-1 plays a role in transporting pri-miR-34a from the nucleus to the cytoplasm. Further studies are necessary to elucidate the mechanism involved in the transport of quiescence-induced pri-miRNAs to the cytoplasm, their (TMG)-capping, and their cytoplasmic processing.

We also discovered the preferential expression of a smaller, alternatively spliced cytoplasmic isoform of Drosha during quiescence (Fig. 6A-E). Recent studies demonstrated the existence of splicing

isoforms of Drosha found in the cytoplasm in several cell lines (48, 49). These Drosha isoforms contain deletions in the nuclear localization region (RS-rich) similar to the small isoform of Drosha found in quiescent HFFs. These data support the hypothesis that splicing of specific genes, such as Drosha, by the spliceosome machinery is important in the regulation of cell cycle progression (54). Importantly, our data expand the concept that a specific Drosha isoform can be found in the cytoplasm of normal human cells in a particular stress condition such as serum starvation. Another study showed that certain RNA viruses trigger Exportin-1-dependent translocation of Drosha to the cytoplasm, providing a potential antiviral mechanism that acts by cleaving viral genomic RNA (55). Consistent with this model, Exportin-1 knock-down reduced the amount of cytoplasmic pri-miR-34a (Fig. 5C), and a short cytoplasmic isoform of Drosha appeared to be involved in pri-miR-34a processing during quiescence (Fig. 6 F and G). Our discovery of pri-miRNAs and a small isoform of Drosha in the cytoplasm suggest the existence of a new miRNA biogenesis pathway during quiescence.

E. The contribution of quiescence-induced miRNAs to cellular growth arrest.

Do the quiescence-induced miRNAs play a role in growth arrest? These miRNAs may be essential for the induction of quiescence, thereby necessitating the use of an alternative miRNA biogenesis pathway when the canonical pathway is repressed. For example, the p53-regulated miR-34a regulates cellular pathways such as cell cycle, differentiation, and apoptosis (56, 57); it is a master regulator of tumor suppression in different human cancers and is being evaluated as a miRNA therapeutic (58). Similarly, miR-638 up-regulation in vascular smooth muscle cells was recently reported to inhibit cell proliferation and migration by targeting the orphan nuclear receptor 1 (NOR1) and consequently affecting the expression of cyclin D1 mRNA (59). In addition, miR-1228* over-expression suppresses xenograft tumor formation by negatively regulating NF-KB activity (60). Finally, miR-3188 was found to reduce cell cycle transition and proliferation in nasopharyngeal carcinoma by targeting mTOR mRNA and regulating PI3K/AKT signaling pathway through FOXO1 (61). In addition, it is possible that impaired processing of specific miRNAs in certain growth conditions may contribute to embryonic lethality of TGS1 knock-out mice (62).

The results of the cell cycle analysis in *SI Appendix*, Fig. S7E show that Exportin-1 regulated miRNAs are not required for entry into senescence. However, the reduced number of cells in S phase 24 h after serum addition suggests that Exportin-1 is required for proper exit from quiescence. We speculate that the Exportin-1-dependent, quiescence-induced miRNAs are important not to maintain growth arrest but rather to poise cells so that they can efficiently reenter the cell cycle when favorable growth conditions are restored.

Not only do miRNAs affect cell growth, but also cell growth state can affect miRNAs. We reported previously that growth arrest can convert some miRNAs from repressors into activators of translation (38), and we show here that quiescence inhibits canonical miRNA biogenesis but stimulates an alternative miRNA biogenesis pathway. Thus, the cellular growth state may globally affect miRNA synthesis and function, which in turn may profoundly influence cellular gene expression and phenotype.

In summary, we have shown that a set of miRNAs is processed independently of the canonical miRNA pathway in quiescent cells via (TMG)-cap modification of their pri-miRNA, interaction with Exportin-1 and cytoplasmic processing by a novel isoform of Drosha (*SI Appendix*, Fig. S16). This pathway is constitutively active in proliferating cells, but its activity is enhanced during quiescence, presumably to regulate key processes involved in cellular growth arrest. Conversely, in quiescent cells the canonical miRNA pathway is turned off, likely contributing to the quiescent state. Future investigation will be necessary to identify additional components of this miRNA Exportin-1-dependent pathway and its impact on cell physiology.

VII. Materials and Methods

Detailed experimental protocols are described in the Supporting Information Appendix (p.175).

A. Cell culture and reagents

Normal primary human foreskin fibroblasts, HeLa/E6-5K, 293T, and C127 cells were cultured in DMEM-10 at 37°C in the presence of 5% CO₂. Further details are in the *SI Appendix*.

B. Quiescence induced by serum starvation and confluency

HFFs, HeLa, and C127 cells were rinsed with DPBS, detached with 0.25% trypsin-EDTA and plated in serum-free DMEM at low density. For quiescence induced by confluency, we plated HFFs (2 X 10⁶ cells/100 mm plate) in DMEM-10. Further details are in the *SI Appendix*.

C. Immunoprecipitation of tri-methylated capped RNA

HFFs were incubated for 72 h in serum-free DMEM to induce quiescence. RNA was extracted using Trizol Reagent and 10 µg of RNA was diluted in NET-2 buffer, pre-cleared and incubated with

Protein G Sepharose 4 Fast Flow beads loaded with 15 μ L normal rabbit serum or 15 μ L rabbit m3G-cap antibody. Beads were rinsed 5 times with NET-2 buffer and resuspended in G-50 buffer. RNA was extracted from the beads by phenol-chloroform-isoamyl alcohol extraction. Further details are in the *SI Appendix*.

VIII. References

1. Pardee AB (1974) A restriction point for control of normal animal cell proliferation. *Proceedings of the National Academy of Sciences of the United States of America* 71(4):1286-1290.
2. O'Farrell PH (2011) Quiescence: early evolutionary origins and universality do not imply uniformity. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 366(1584):3498-3507.
3. Collier HA (2011) Cell biology. The essence of quiescence. *Science* 334(6059):1074-1075.
4. Valcourt JR, *et al.* (2012) Staying alive: metabolic adaptations to quiescence. *Cell cycle* 11(9):1680-1696.
5. Collier HA, Sang L, & Roberts JM (2006) A new description of cellular quiescence. *PLoS biology* 4(3):e83.
6. Cheung TH & Rando TA (2013) Molecular regulation of stem cell quiescence. *Nature reviews. Molecular cell biology* 14(6):329-340.
7. Bohnsack MT, Czaplinski K, & Gorlich D (2004) Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *Rna* 10(2):185-191.
8. Krol J, Loedige I, & Filipowicz W (2010) The widespread regulation of microRNA biogenesis, function and decay. *Nature reviews. Genetics* 11(9):597-610.
9. Lee Y, Han J, Yeom KH, Jin H, & Kim VN (2006) Drosha in primary microRNA processing. *Cold Spring Harbor symposia on quantitative biology* 71:51-57.
10. Lee Y, *et al.* (2003) The nuclear RNase III Drosha initiates microRNA processing. *Nature* 425(6956):415-419.
11. Gregory RI, *et al.* (2004) The Microprocessor complex mediates the genesis of microRNAs. *Nature* 432(7014):235-240.
12. Denli AM, Tops BB, Plasterk RH, Ketting RF, & Hannon GJ (2004) Processing of primary microRNAs by the Microprocessor complex. *Nature* 432(7014):231-235.
13. Yi R, Qin Y, Macara IG, & Cullen BR (2003) Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & development* 17(24):3011-3016.
14. Hutvagner G, *et al.* (2001) A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* 293(5531):834-838.
15. Grishok A, *et al.* (2001) Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* 106(1):23-34.
16. Knight SW & Bass BL (2001) A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* 293(5538):2269-2271.
17. Han J, *et al.* (2004) The Drosha-DGCR8 complex in primary microRNA processing. *Genes & development* 18(24):3016-3027.
18. Lian SL, *et al.* (2009) The C-terminal half of human Ago2 binds to multiple GW-rich regions of GW182 and requires GW182 to mediate silencing. *Rna* 15(5):804-813.
19. Hammond SM, Boettcher S, Caudy AA, Kobayashi R, & Hannon GJ (2001) Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* 293(5532):1146-1150.
20. Mourelatos Z, *et al.* (2002) miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes & development* 16(6):720-728.
21. Ruby JG, Jan CH, & Bartel DP (2007) Intronic microRNA precursors that bypass Drosha processing. *Nature* 448(7149):83-86.

22. Cazalla D, Xie M, & Steitz JA (2011) A primate herpesvirus uses the integrator complex to generate viral microRNAs. *Molecular cell* 43(6):982-992.
23. Havens MA, Reich AA, Duelli DM, & Hastings ML (2012) Biogenesis of mammalian microRNAs by a non-canonical processing pathway. *Nucleic acids research* 40(10):4626-4640.
24. Cifuentes D, *et al.* (2010) A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* 328(5986):1694-1698.
25. Cheloufi S, Dos Santos CO, Chong MM, & Hannon GJ (2010) A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* 465(7298):584-589.
26. Shapiro JS, Langlois RA, Pham AM, & Tenoever BR (2012) Evidence for a cytoplasmic microprocessor of pri-miRNAs. *Rna* 18(7):1338-1346.
27. Xie M, *et al.* (2013) Mammalian 5'-capped microRNA precursors that generate a single microRNA. *Cell* 155(7):1568-1580.
28. Hutten S & Kehlenbach RH (2007) CRM1-mediated nuclear export: to the pore and beyond. *Trends in cell biology* 17(4):193-201.
29. Ohno M, Segref A, Bachi A, Wilm M, & Mattaj IW (2000) PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* 101(2):187-198.
30. Palacios I, Hetzer M, Adam SA, & Mattaj IW (1997) Nuclear import of U snRNPs requires importin beta. *The EMBO journal* 16(22):6783-6792.
31. Kiss T (2004) Biogenesis of small nuclear RNPs. *Journal of cell science* 117(Pt 25):5949-5951.
32. Boulon S, *et al.* (2004) PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. *Molecular cell* 16(5):777-787.
33. Yedavalli VS & Jeang KT (2010) Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs. *Proceedings of the National Academy of Sciences of the United States of America* 107(33):14787-14792.
34. Medina R, *et al.* (2008) MicroRNAs 221 and 222 bypass quiescence and compromise cell survival. *Cancer research* 68(8):2773-2780.
35. Suh EJ, *et al.* (2012) A microRNA network regulates proliferative timing and extracellular matrix synthesis during cellular quiescence in fibroblasts. *Genome biology* 13(12):R121.
36. Martinez I, Cazalla D, Almstead LL, Steitz JA, & DiMaio D (2011) miR-29 and miR-30 regulate B-Myb expression during cellular senescence. *Proceedings of the National Academy of Sciences of the United States of America* 108(2):522-527.
37. Vasudevan S, Tong Y, & Steitz JA (2007) Switching from repression to activation: microRNAs can up-regulate translation. *Science* 318(5858):1931-1934.
38. Iwasaki YW, *et al.* (2013) Global microRNA elevation by inducible Exportin 5 regulates cell cycle entry. *Rna* 19(4):490-497.
39. Link S, Grund SE, & Diederichs S (2016) Alternative splicing affects the subcellular localization of Drosha. *Nucleic acids research* 44(11):5330-5343.
40. Rissland OS, Hong SJ, & Bartel DP (2011) MicroRNA destabilization enables dynamic regulation of the miR-16 family in response to cell-cycle changes. *Molecular cell* 43(6):993-1004.
41. Bussing I, Yang JS, Lai EC, & Grosshans H (2010) The nuclear export receptor XPO-1 supports primary miRNA processing in *C. elegans* and *Drosophila*. *The EMBO journal* 29(11):1830-1839.
42. Castanotto D, Lingeman R, Riggs AD, & Rossi JJ (2009) CRM1 mediates nuclear-cytoplasmic shuttling of mature microRNAs. *Proceedings of the National Academy of Sciences of the United States of America* 106(51):21655-21659.
43. Luhrmann R, *et al.* (1982) Isolation and characterization of rabbit anti-m3 2,2,7G antibodies. *Nucleic acids research* 10(22):7103-7113.
44. Girard C, *et al.* (2008) Characterization of a short isoform of human Tgs1 hypermethylase associating with small nucleolar ribonucleoprotein core proteins and produced by limited proteolytic processing. *The Journal of biological chemistry* 283(4):2060-2069.
45. Chong MM, *et al.* (2010) Canonical and alternate functions of the microRNA biogenesis machinery. *Genes & development* 24(17):1951-1960.

46. Kawai S & Amano A (2012) BRCA1 regulates microRNA biogenesis via the DROSHA microprocessor complex. *The Journal of cell biology* 197(2):201-208.
47. Kim YK, Kim B, & Kim VN (2016) Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis. *Proceedings of the National Academy of Sciences of the United States of America* 113(13):E1881-1889.
48. Ha M & Kim VN (2014) Regulation of microRNA biogenesis. *Nature reviews. Molecular cell biology* 15(8):509-524.
49. Legesse-Miller A, *et al.* (2012) Quiescent fibroblasts are protected from proteasome inhibition-mediated toxicity. *Molecular biology of the cell* 23(18):3566-3581.
50. Gibbings D, *et al.* (2012) Selective autophagy degrades DICER and AGO2 and regulates miRNA activity. *Nature cell biology* 14(12):1314-1321.
51. Bitto A, *et al.* (2010) Long-term IGF-I exposure decreases autophagy and cell viability. *PloS one* 5(9):e12592.
52. Verheggen C & Bertrand E (2012) CRM1 plays a nuclear role in transporting snoRNPs to nucleoli in higher eukaryotes. *Nucleus* 3(2):132-137.
53. Karamysheva Z, Diaz-Martinez LA, Warrington R, & Yu H (2015) Graded requirement for the spliceosome in cell cycle progression. *Cell cycle* 14(12):1873-1883.
54. Shapiro JS, *et al.* (2014) Drosha as an interferon-independent antiviral factor. *Proceedings of the National Academy of Sciences of the United States of America* 111(19):7108-7113.
55. He L, *et al.* (2007) A microRNA component of the p53 tumour suppressor network. *Nature* 447(7148):1130-1134.
56. Chen F & Hu SJ (2012) Effect of microRNA-34a in cell cycle, differentiation, and apoptosis: a review. *Journal of biochemical and molecular toxicology* 26(2):79-86.
57. Bader AG (2012) miR-34 - a microRNA replacement therapy is headed to the clinic. *Frontiers in genetics* 3:120.
58. Li P, *et al.* (2013) MicroRNA-638 is highly expressed in human vascular smooth muscle cells and inhibits PDGF-BB-induced cell proliferation and migration through targeting orphan nuclear receptor NOR1. *Cardiovascular research* 99(1):185-193.
59. Jia L, *et al.* (2013) Restoration of miR-1228* expression suppresses epithelial-mesenchymal transition in gastric cancer. *PloS one* 8(3):e58637.
60. Lund E, Guttinger S, Calado A, Dahlberg JE, & Kutay U (2004) Nuclear export of microRNA precursors. *Science* 303(5654):95-98.
61. Jia Y, *et al.* (2012) Early embryonic lethality of mice with disrupted transcription cofactor PIMT/NCOA6IP/Tgs1 gene. *Mechanisms of development* 129(9-12):193-207.

IX. Figure Legends

Figure 1. Differential expression of mature miRNAs and their corresponding pri-miRNAs during quiescence. (A) Heat map of mature miRNA microarray data comparing proliferating HFFs to HFFs serum-starved (SS) for 24, 48, or 72 h. Only differentially expressed miRNAs during quiescence are shown. (B) qRT-PCR validation of representative differentially expressed mature miRNAs during quiescence. (C) qRT-PCR quantitation of pri-miRNAs corresponding to differentially expressed mature miRNAs (using primers that bind upstream and downstream of the miRNA stem loop). Red bars represent the miRNAs up-regulated and the green bars represent the miRNAs down-regulated during quiescence by miRNA microarray analysis. Numbers refer to days of serum starvation. Black bars represent miRNAs in proliferating (P) cells. RNU43 and GAPDH mRNA were used to normalize the qRT-PCR analyses, which are shown relative to expression in proliferating cells. Similar results were obtained in three independent experiments.

Figure 2. Regulation of Exportin-5 during quiescence by autophagy and miR-34a. (A) Western blot showing reduction of Exportin-5 but not Exportin-1 levels during the induction of quiescence by serum starvation. p27 was used as a marker of quiescence induction, and β -actin was used as a loading control. (B) Western blot showing Exportin-5 and MDM2 in proliferating and quiescent HFFs treated with 10 μ M of the proteasome inhibitor, MG132. Mock indicates cells proliferating in the presence of serum; cells were serum-starved for 3, 6, or 12 h as indicated to induce quiescence. (C) Western blot of Exportin-5 and p62 in proliferating and quiescent HFFs treated with 10 μ M of the autophagy inhibitor, 3-MA. Similar results were obtained in three independent experiments. (D) The sequence of hsa-miR-34a and its putative wild-type (WT) or mutant (mt) binding site in the Exportin-5 3'UTR are shown. Vertical lines indicate predicted base pairing between hsa-miR-34a and the Exportin-5 3' UTR, and the dots indicate potential G-U base pairs. Point mutations in the Exportin-5 3'UTR are shown with asterisks. Numbers represent the nucleotide sequence of Exportin-5 3'UTR numbered from the beginning of the untranslated region. (E) miR-34a regulates the expression of Exportin-5. Left panels, Western blot of Exportin-5 and Exportin-1 48 h after transfection of exogenous miR-34a. Right panel, qRT-PCR analysis of Exportin-5 mRNA 48 h after transfection of miR-34a. (F) Activity of the Exportin-5 3'UTR is regulated by cell growth state. Luciferase activity from the empty reporter vector (Control), the vector containing the wild-type Exportin-5 3'UTR (WT) or the Exportin-5 3'UTR containing a mutant miR-34a binding site (mt). Assays were conducted in proliferating (P) or quiescent (Q) HFFs as described in the experimental procedures. (G) miR-34a regulates the Exportin-5 3'UTR. Luciferase activity from the reporter vectors described in panel (F) was measured in proliferating HFFs transfected with a scrambled miRNA control (miR-CTRL) or miR-34a. In panels (F) and (G) firefly luciferase activity was measured and normalized to Renilla luciferase activity expressed from the same plasmid (Relative luciferase units, RLUs). Data were obtained in triplicate, and the graph shows the average of two individual experiments. Two-tailed *t*-test results are indicated by * for $P < 0.05$ and ** for $P < 0.01$.

Figure 3. Biogenesis of quiescence-induced miRNAs requires Exportin-1 but not Exportin-5. (A) qRT-PCR analysis showing relative expression of primary miRNAs (light gray bars) and their corresponding mature miRNAs (dark gray bars) in proliferating HFFs 48 h after transfection with an siRNA scrambled control (dashed line) or an siRNA targeting Exportin-5 (Dharmacon) (dark and light gray bars). Right panel shows Western blot analysis of Exportin-1 and Exportin-5 after siRNA transfection. (B) qRT-PCR analysis showing relative levels of primary miRNAs (light gray bars) and their corresponding mature miRNAs (dark gray bars) in proliferating HFFs 48 h after transfection with an siRNA scrambled control (dashed line) or an siRNA targeting Exportin-1 (ID s14937, Ambion) (dark and light gray bars). Two-tailed *t*-test results are indicated by * for $P < 0.05$ and ** for $P < 0.01$. Western blot analysis of Exportin-1 and Exportin-5 is also shown after siRNA transfection. (C) qRT-PCR analysis of quiescence-induced miRNAs and miR-34a Northern blot 48 h after transfection of an siRNA scrambled control (black bars) or an siRNA targeting Exportin-1 (ID s14937, Ambion) (gray bars) into proliferating HFFs in the presence of serum (P) or into

HFFs serum-starved (SS) for 24 or 72 h (24, 48, or 72 h for miR-34a Northern blot). Transfection of quiescent cells was done 24 h prior to serum starvation. (D) Exportin-1 physically associates with pri-miR-34a in quiescent cells. qRT-PCR were used to detect U3 snoRNA, pri-miR-34a, or pri-miR-423 in control Flag-Vector (Flag, black bars) or Flag-Exportin-1 (Flag-XPO1, gray bars) immunoprecipitates from quiescent HFFs (expression data in Fig. S7E). Exogenous spike of *C. elegans* RNA was used for qRT-PCR normalization by amplifying Ams-1 mRNA. Similar results were obtained in two independent experiments.

Figure 4. Quiescence-induced primary miRNAs are (TMG)-capped by TGS1. (A) Quiescence-induced pri-miRNAs contain a TMG cap. Extracts were prepared from proliferating or serum starved quiescent HFFs transfected with control siRNA or siRNA targeting TGS1 (ID s41313, Ambion). siRNAs were transfected 24 h prior to serum starvation. RNA was harvested 48 h after transfection of siRNAs in proliferating HFFs or 72 h after removal of serum. Extracts were then immunoprecipitated with control antibody or antibody recognizing the TMG cap, and immunoprecipitated U3 snoRNA and pri-miRNAs were amplified by RT-PCR and visualized after gel electrophoresis. Labels: input (I), supernatant (S), immunoprecipitate (IP), control anti-rabbit serum (C), and anti-TMG antibody (T). (B) TGS1 knock-down inhibits expression of quiescence-induced miRNAs. HFFs were transfected with an siRNA scrambled control (black bars) or an siRNA targeting TGS1 (sc-45875, Santa Cruz) (gray bars). 24 h later, cells were serum starved (SS) for 24 or 72 h, or maintained in serum (P). RNA was then prepared and quantified by qRT-PCR. (C) TGS1 knock-down affects the binding of pri-miR-34a to Exportin-1 in quiescent cells. qRT-PCR was used to detect U3 snoRNA, pri-miR-34a, or pri-miR-423 in immunoprecipitate samples of control Flag-Vector or Flag-Exportin-1 in quiescent HFF cells. Cells were transfected with siRNA control (black bars) or siRNA against TGS1 (gray bars). Two-tailed t-test results are indicated by ** for $P < 0.01$. Similar results were obtained in two independent experiments.

Figure 5. Detection of cytoplasmic pri-miR-34a is dependent on Exportin-1 expression. (A) Detection of pri-miR-34a in the cytoplasmic fraction of Drosha knock-down HFFs. Proliferating (P) and quiescent (Q) HFFs were transfected with control siRNA or siRNA targeting Drosha (sc-44080, Santa Cruz) and after 48 h subjected to cytoplasmic and nuclear RNA fractionation. RT-PCR was used to detect pri-miR-34a, pri-miR-29b, and pri-miR-423. U6 snRNA is a nuclear control and mitochondrial (mt) tRNA-val is a cytoplasmic marker. (B) qRT-PCR of cytoplasmic fractions as in Figure 5A was used to quantify the levels of pri-miR-34a. Two-tailed *t*-test results are indicated by ** for $P < 0.01$. (C and D) Exportin-1 knock-down affects the distribution of pri-miR-34a in quiescent HFFs. Cells were transfected with an siRNA scrambled control, an siRNA targeting Drosha (sc-44080, Santa Cruz), or siRNAs targeting Drosha (sc-44080, Santa Cruz) and Exportin-1 (ID s14937, Ambion). 24 h later, cells were serum starved for 72 h. RNA was then prepared from the cytoplasmic (C) or nuclear (D) fraction for the quantification of pri-miR-34a and pri-miR-423 by qRT-PCR. Normalization was done using *C. elegans* total RNA as an exogenous spike for the amplification of the worm-specific *ama-1* gene. Similar results were obtained in two independent experiments.

Figure 6. Detection of a small cytoplasmic isoform of Drosha during quiescence. (A) Western blot showing reduction of full-length Drosha and appearance of a more rapidly migrating form of Drosha (Drosha_{short}) during the induction of quiescence by serum starvation. (B) A small form of Drosha (Drosha_{short}) is found in the cytoplasmic fraction of quiescent cells. Western blot analysis of Drosha and DGCR8 in total, nuclear, and cytoplasmic fractions in proliferating (P) and quiescent (Q) HFFs. 10% of the total loading was used for the nuclear fraction in the Drosha Western blot. Histone deacetylase 1 (HDAC1) was used as nuclear marker and β -tubulin as cytoplasmic marker. (C) Cytoplasmic localization of Drosha in quiescent HFF cells. Immunofluorescence of Drosha (green) in proliferating and quiescent HFFs. Nuclear fluorescence (DAPI, blue) and cytoskeleton immunofluorescence (actin, red) were used to visualize the localization of Drosha. (D) Quantification of cytoplasmic Drosha immunofluorescence. Proliferating and quiescent HFF cells were incubated with digitonin (to selectively permeabilize only the plasma membrane) to diminish masking of cytoplasmic Drosha staining with nuclear Drosha signal. Confocal microscopy was used to

visualize the cells, and fluorescent images were used to obtain relative pixel area quantification using Image J software. Two-tailed *t*-test results are indicated by * for $P < 0.05$. Similar results were obtained in three independent experiments. (E) Drosha mRNA exon skipping during quiescence. Illustration and sequencing data show two alternative splicing events (red lines) between exon 3 and 4 and between exon 5 and 7 in Drosha mRNA and the affected regions in Drosha protein. Numbers represent the nucleotide sequence of Drosha mRNA. P-rich indicates the proline-rich region, RS-rich indicates the arginine and serine-rich region, RIIIDa and RIIIDb indicate RNase III domains a and b, respectively, and dsRBD indicates double-stranded RNA-binding domain. (F) The short isoform of Drosha is responsible for processing of pri-miR-34a. Quantitative RT-PCR of pri-miR-34a and pri-miR-423 was conducted in quiescent HFFs transfected with control siRNA or siRNA against the full-length Drosha (IDs26491, Ambion) or both Drosha isoforms (both) (sc-44080, Santa Cruz). Reduced accumulation of pri-miR-34a in cells treated with siRNA targeting full-length Drosha compared to siRNA against both Drosha isoforms (both) suggests the short Drosha isoform can process pri-miR-34a. (G) qRT-PCR of mature miR-34a and miR-423-3p was conducted from the same experiments shown in Figure 6F. Two-tailed *t*-test results are indicated by ** for $P < 0.01$. Similar results were obtained in two independent experiments.

X. Figures

Figure 1

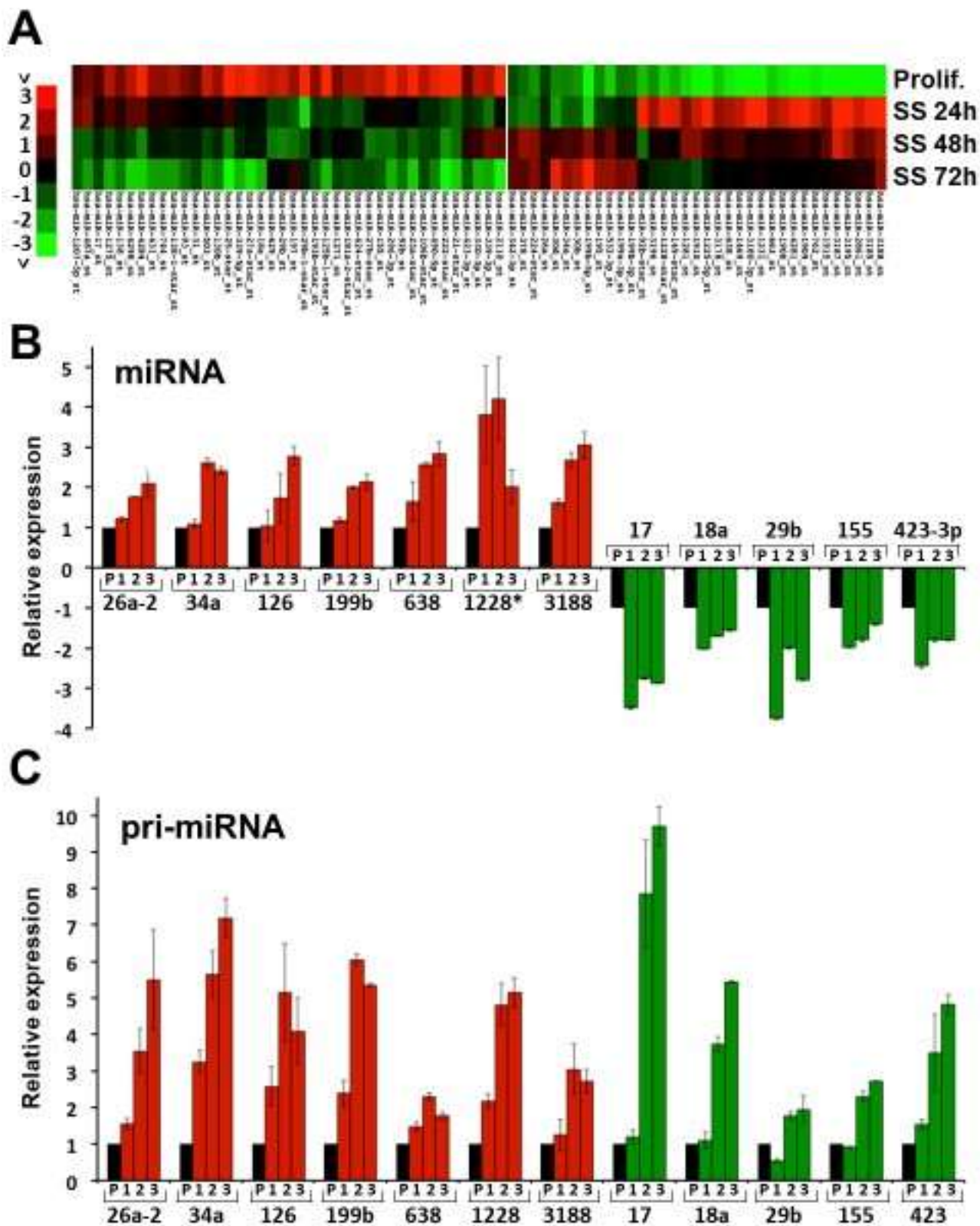


Figure 2

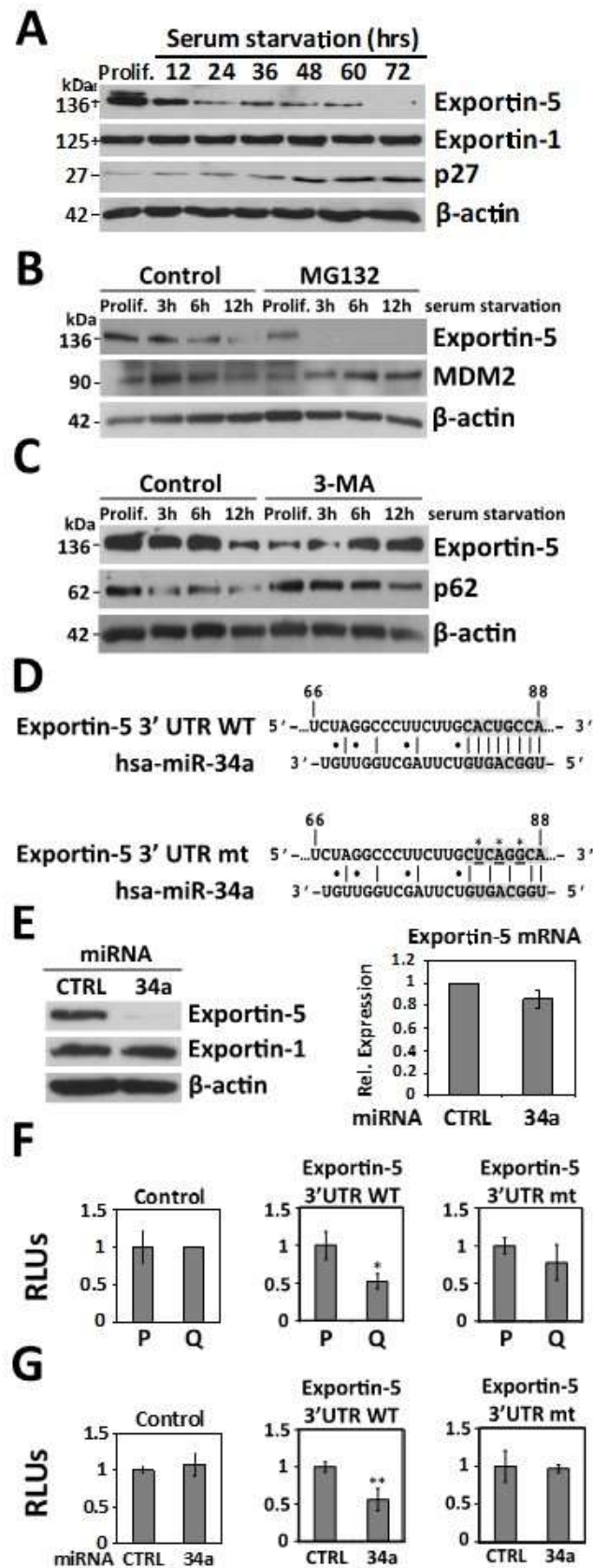


Figure 3

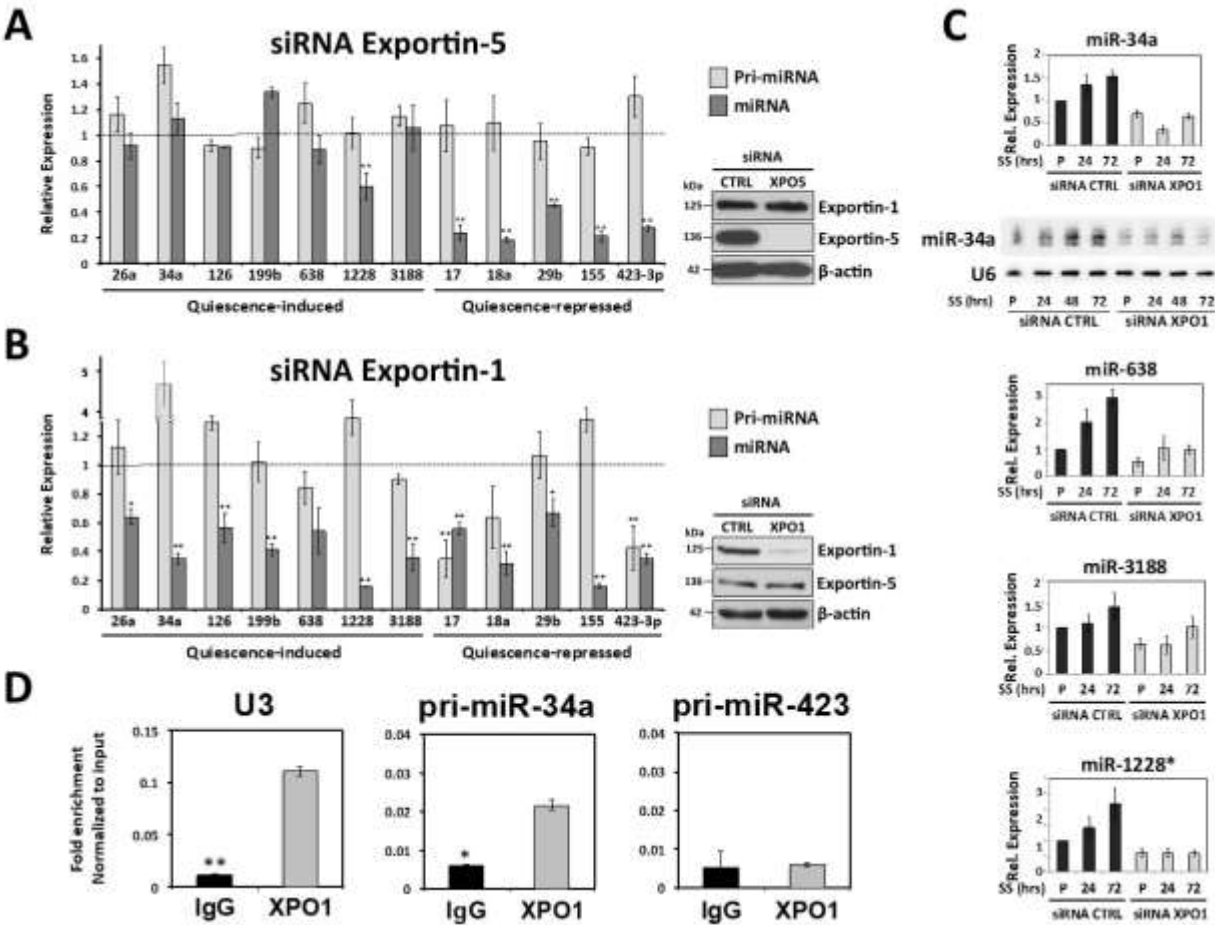


Figure 4

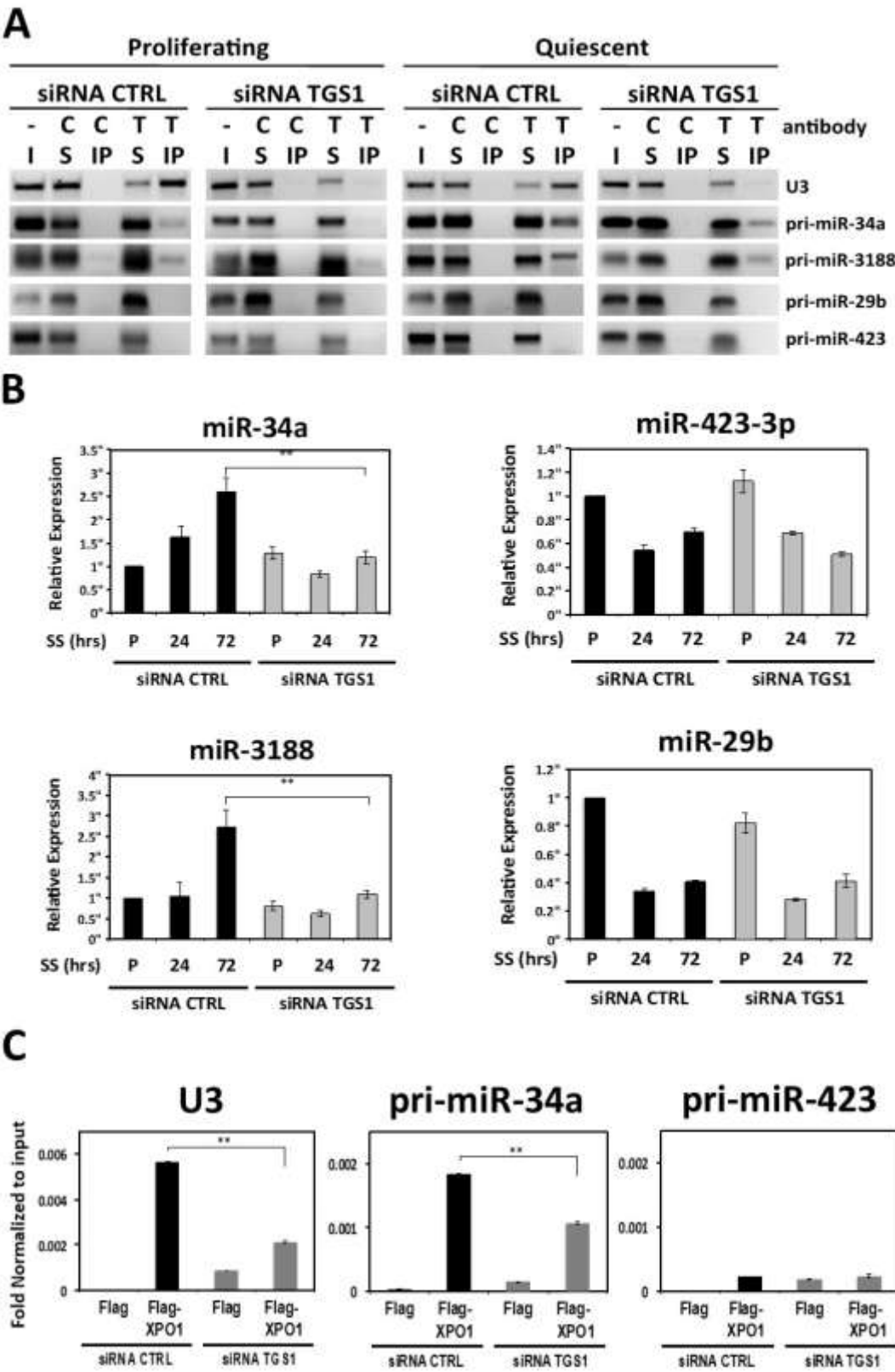


Figure 5

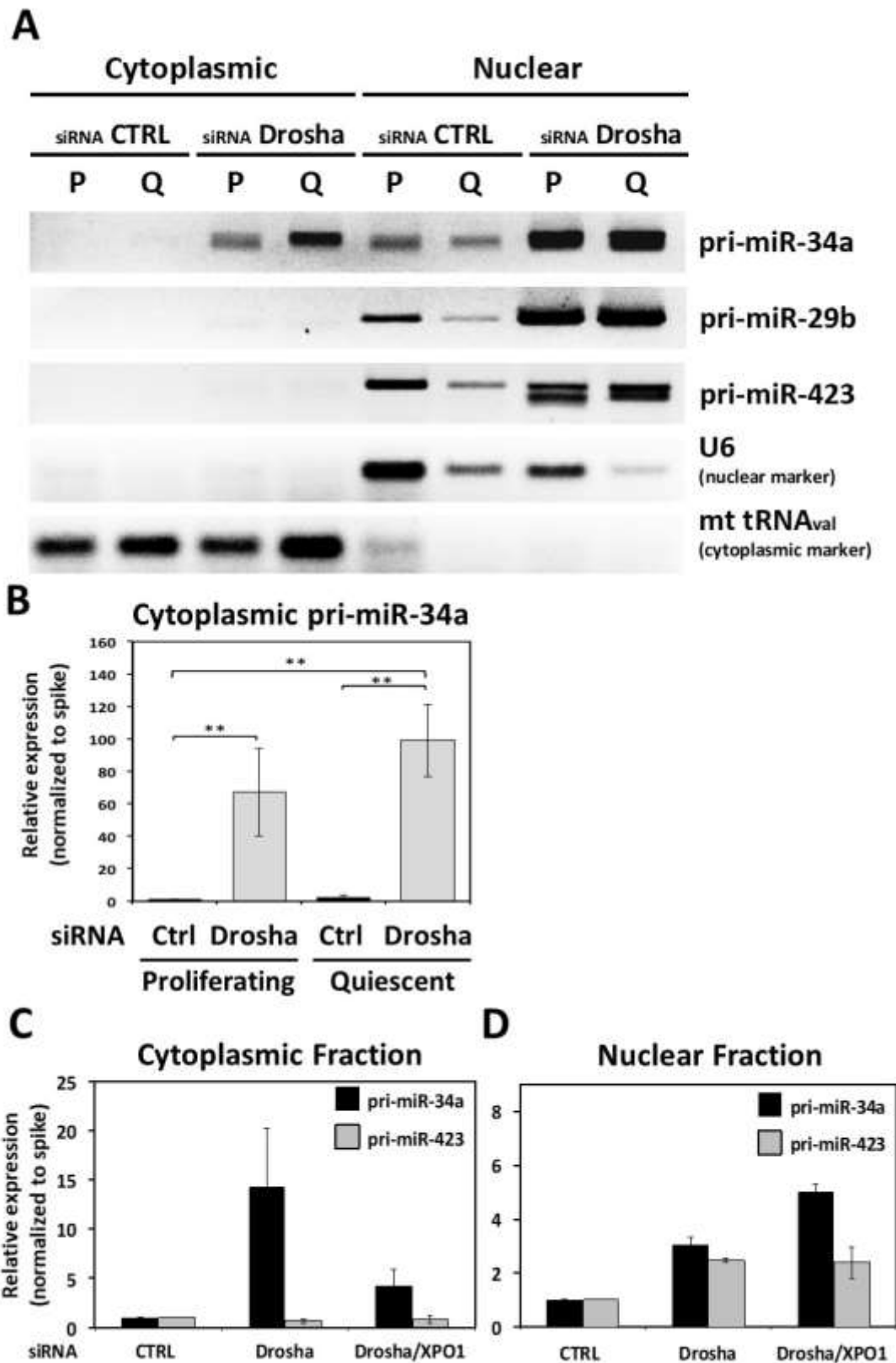
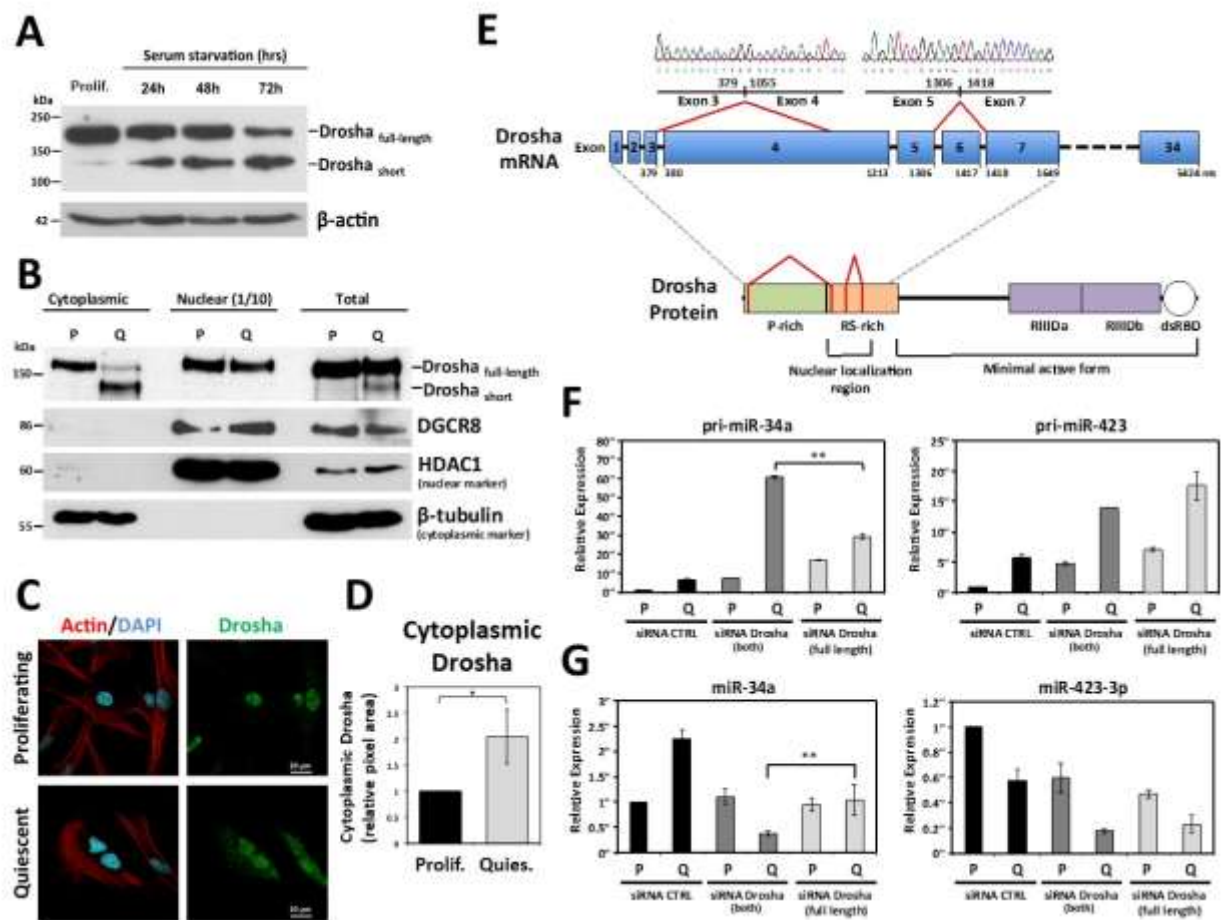


Figure 6



XI. Supplementary Tables

Table S1

miRNA Name	12hrs SS vs mock	48hrs SS vs mock	72hrs SS vs mock
hsa-let-7g-star_st	1.15	1.53	2.23
hsa-miR-101_st	1.68	1.79	1.94
hsa-miR-1208_st	2.10	1.55	1.46
hsa-miR-1225-5p_st	3.21	2.13	1.70
hsa-miR-1228-star_st	4.48	2.10	1.67
hsa-miR-1231_st	5.31	2.61	2.47
hsa-miR-1246_st	3.00	3.68	4.87
hsa-miR-126_st	1.18	1.79	3.16
hsa-miR-1281_st	2.04	2.05	1.37
hsa-miR-133b_st	1.71	1.55	1.87
hsa-miR-1469_st	5.39	3.37	2.54
hsa-miR-149-star_st	3.90	1.78	1.46
hsa-miR-150-star_st	2.62	1.90	1.53
hsa-miR-154_st	1.39	1.71	2.74
hsa-miR-1825_st	1.82	1.69	1.25
hsa-miR-187-star_st	2.92	3.97	1.76
hsa-miR-1908_st	4.59	2.49	2.21
hsa-miR-1909_st	5.49	2.72	2.32
hsa-miR-1909-star_st	1.92	1.92	1.50
hsa-miR-1910_st	2.73	2.36	1.39
hsa-miR-1915_st	4.59	3.64	2.58
hsa-miR-194_st	2.04	2.05	3.50
hsa-miR-195_st	1.15	1.50	1.83
hsa-miR-199a-3p_st	1.34	1.66	2.33
hsa-miR-199b-3p_st	1.35	1.78	2.54
hsa-miR-199b-5p_st	1.21	2.42	4.31
hsa-miR-224-star_st	1.05	1.74	1.71
hsa-miR-2277_st	3.10	1.95	1.34
hsa-miR-26a_st	1.09	1.55	1.57
hsa-miR-26b_st	1.86	1.38	2.18
hsa-miR-2861_st	7.14	4.60	3.70
hsa-miR-2909_st	1.63	1.67	1.32
hsa-miR-301a_st	1.78	1.74	1.66
hsa-miR-302a_st	1.52	1.56	1.53
hsa-miR-30b_st	1.13	1.64	2.13

hsa-miR-30d_st	-1.05	1.54	2.34
hsa-miR-30e-star_st	1.77	1.50	2.00
hsa-miR-3144-5p_st	2.09	1.68	1.39
hsa-miR-3154_st	1.65	1.32	1.51
hsa-miR-3168_st	1.55	1.66	1.78
hsa-miR-3178_st	3.79	2.50	2.12
hsa-miR-3180-3p_st	4.44	2.24	2.13
hsa-miR-3181_st	4.08	3.60	3.01
hsa-miR-3185_st	7.28	4.50	3.59
hsa-miR-3187_st	5.49	4.27	2.83
hsa-miR-3188_st	8.77	5.21	4.93
hsa-miR-3194_st	1.12	2.15	1.59
hsa-miR-3195_st	6.20	4.79	3.22
hsa-miR-3196_st	3.29	1.87	1.48
hsa-miR-342-3p_st	-1.05	1.55	1.61
hsa-miR-346_st	3.20	2.69	2.15
hsa-miR-34a_st	1.12	1.95	2.39
hsa-miR-34a-star_st	1.01	1.55	2.44
hsa-miR-378_st	-1.04	1.82	1.58
hsa-miR-422a_st	1.81	2.87	3.07
hsa-miR-4259_st	1.64	1.49	1.61
hsa-miR-4271_st	2.69	1.98	1.56
hsa-miR-4281_st	5.29	2.70	2.32
hsa-miR-4292_st	2.18	2.06	1.90
hsa-miR-4320_st	1.58	1.67	1.59
hsa-miR-4321_st	2.18	1.90	1.31
hsa-miR-4322_st	3.43	2.57	2.47
hsa-miR-455-5p_st	2.41	2.43	3.06
hsa-miR-498_st	2.39	1.71	1.20
hsa-miR-514_st	1.30	1.56	1.50
hsa-miR-523_st	-1.19	1.68	1.55
hsa-miR-532-3p_st	1.33	1.81	1.99
hsa-miR-557_st	1.89	1.22	1.54
hsa-miR-564_st	2.44	2.30	1.79
hsa-miR-566_st	1.70	1.25	1.55
hsa-miR-572_st	3.60	2.18	1.21
hsa-miR-575_st	1.57	1.66	1.53
hsa-miR-596_st	2.92	1.80	1.08
hsa-miR-602_st	4.50	1.90	2.12
hsa-miR-638_st	4.67	3.06	2.30
hsa-miR-654-5p_st	1.69	2.12	1.33

hsa-miR-663_st	3.62	2.22	1.95
hsa-miR-664-star_st	1.29	1.85	1.87
hsa-miR-675_st	1.59	1.52	1.67
hsa-miR-675-star_st	1.76	1.56	1.71
hsa-miR-762_st	4.73	2.52	2.17
hsa-miR-768-3p_st	1.11	1.67	2.78
hsa-miR-885-3p_st	3.26	2.56	2.34
hsa-miR-885-5p_st	2.61	2.17	1.81
hsa-miR-888-star_st	1.61	1.25	1.52
hsa-miR-92b-star_st	3.45	1.19	1.61
hsa-miR-943_st	3.41	1.69	2.13
hsa-miR-99a-star_st	1.41	1.85	1.59

Table S1. Mature microRNAs up-regulated during the induction of quiescence in HFFs. Table show microRNAs that were up-regulated 1.5-fold or greater in at least two-time points of serum starvation (red values) in comparison to proliferating HFF cells. GeneChip® miRNA 2.0 Array (Affymetrix, Santa Clara, CA) were used for this analysis.

Table S2

miRNA Name	12hrs SS vs mock	48hrs SS vs mock	72hrs SS vs mock
hsa-let-7a-2-star_st	-1.92	-1.39	-1.58
hsa-miR-106b-star_st	-1.71	-1.84	-1.84
hsa-miR-1184_st	-1.17	-2.08	-1.62
hsa-miR-1207-5p_st	1.08	-1.59	-1.59
hsa-miR-1244_st	-1.50	-1.51	1.30
hsa-miR-1247_st	-1.95	-1.80	-1.17
hsa-miR-1249_st	-1.58	-1.49	-1.74
hsa-miR-125b-1-star_st	-2.50	-2.33	-3.34
hsa-miR-1271_st	-1.54	-1.35	-1.59
hsa-miR-1275_st	-1.30	-1.95	-2.44
hsa-miR-1285_st	-1.20	-1.89	-2.51
hsa-miR-1292_st	1.35	-1.59	-2.87
hsa-miR-1296_st	-2.18	-1.73	-2.20
hsa-miR-1308_st	-1.94	-2.64	-3.48
hsa-miR-130b_st	-1.49	-1.76	-2.24
hsa-miR-138_st	-1.33	-1.78	-2.34
hsa-miR-138-1-star_st	-1.46	-1.72	-2.31
hsa-miR-139-3p_st	-2.04	-3.03	-2.42
hsa-miR-1470_st	-7.07	-7.07	-7.07
hsa-miR-1471_st	-4.02	-1.98	-1.25
hsa-miR-1538_st	-2.36	-2.61	-1.67
hsa-miR-155_st	-1.79	-2.51	-2.42
hsa-miR-155-star_st	-1.86	-1.76	-1.63
hsa-miR-17_st	-1.20	-1.50	-1.69
hsa-miR-181a-2-star_st	-1.78	-1.46	-1.77
hsa-miR-184_st	-1.89	-2.01	-1.62
hsa-miR-185-star_st	-1.84	-1.42	-1.51
hsa-miR-188-5p_st	-1.85	-1.58	-1.28
hsa-miR-18a_st	-1.87	-1.90	-2.81
hsa-miR-18a-star_st	-1.81	-2.26	-1.57
hsa-miR-18b_st	-1.25	-1.77	-1.50
hsa-miR-190_st	-1.28	-1.70	-1.57
hsa-miR-1914_st	-4.22	-3.43	-1.18
hsa-miR-1915-star_st	-1.63	-2.08	-1.55
hsa-miR-193b-star_st	-1.54	-1.34	-1.66
hsa-miR-197_st	-1.72	-1.79	-1.25
hsa-miR-1972_st	-1.74	-1.93	-1.95

hsa-miR-21-star_st	-2.91	-2.96	-3.50
hsa-miR-2110_st	-1.72	-1.37	-3.06
hsa-miR-22-star_st	-1.46	-2.50	-1.67
hsa-miR-222-star_st	-3.11	-3.56	-4.22
hsa-miR-23a-star_st	-2.05	-2.55	-3.15
hsa-miR-25-star_st	-1.97	-2.74	-3.91
hsa-miR-27a-star_st	-1.81	-1.84	-2.85
hsa-miR-27b-star_st	-1.48	-1.81	-1.67
hsa-miR-296-3p_st	-1.66	-2.11	-2.38
hsa-miR-298_st	-1.74	-1.23	-1.50
hsa-miR-29b_st	-1.69	-1.89	-1.52
hsa-miR-29b-1-star_st	-4.04	-2.91	-2.55
hsa-miR-31_st	-1.42	-1.60	-1.89
hsa-miR-3130-5p_st	-2.00	-1.14	-2.56
hsa-miR-3131_st	-2.08	-1.83	-1.84
hsa-miR-3136_st	-1.91	-2.01	-2.19
hsa-miR-3158_st	-1.60	-1.81	-1.39
hsa-miR-3170_st	-1.70	1.08	-1.93
hsa-miR-3183_st	-2.05	-2.32	-1.32
hsa-miR-3198_st	-1.18	-1.56	-1.93
hsa-miR-323-5p_st	-2.64	-1.93	1.25
hsa-miR-330-3p_st	-1.68	-1.31	-2.09
hsa-miR-339-3p_st	-1.75	-1.12	-1.91
hsa-miR-339-5p_st	-1.79	-1.87	-2.54
hsa-miR-33b-star_st	-1.25	-1.61	-2.71
hsa-miR-376a_st	-1.08	-2.53	-1.78
hsa-miR-378b_st	-1.35	-1.51	-1.60
hsa-miR-423-3p_st	-1.51	-1.18	-1.70
hsa-miR-424-star_st	-2.07	-1.75	-2.03
hsa-miR-4284_st	-1.48	-2.51	-2.91
hsa-miR-4298_st	-1.31	-2.18	-3.29
hsa-miR-431_st	-1.31	-1.51	-2.04
hsa-miR-4315_st	-1.17	-1.53	-1.63
hsa-miR-487a_st	1.06	-1.56	-1.75
hsa-miR-490-3p_st	-3.55	-1.88	-1.68
hsa-miR-490-5p_st	-1.92	-2.03	-2.34
hsa-miR-503_st	-1.72	-2.19	-2.71
hsa-miR-508-5p_st	-2.34	-1.49	-1.68
hsa-miR-519c-3p_st	-1.14	-1.59	-1.55
hsa-miR-548a-3p_st	-2.24	-1.20	-1.68
hsa-miR-550_st	-1.58	-2.01	-1.86

hsa-miR-550-star_st	-1.77	-1.85	-1.58
hsa-miR-608_st	-5.90	-2.98	-4.15
hsa-miR-629_st	-1.65	-1.81	-1.44
hsa-miR-629-star_st	-2.25	-2.62	-2.05
hsa-miR-639_st	-1.52	-1.47	-1.65
hsa-miR-658_st	-2.66	-1.88	-1.06
hsa-miR-671-3p_st	-2.03	-2.37	-3.83
hsa-miR-720_st	-1.05	-1.83	-1.56
hsa-miR-744_st	-1.34	-1.59	-2.22
hsa-miR-758_st	-1.73	-1.29	-1.50
hsa-miR-770-5p_st	-1.51	-2.69	-2.08
hsa-miR-886-3p_st	-2.39	-6.22	-5.80
hsa-miR-886-5p_st	-1.84	-2.32	-3.08
hsa-miR-92a-1-star_st	-1.71	-1.73	-1.60
hsa-miR-92b_st	-1.68	-1.90	-2.14
hsa-miR-93_st	-1.50	-1.69	-2.18

Table S2. Mature microRNAs down-regulated during the induction of quiescence in HFFs. Table shows microRNAs that were down-regulated as defined in the legend to Table S1, shown in green.

Table S3

primers	sequence (5'-3')
pri-miR-26a-2 FWD	ACCTAATCATGACCTGGACAGAC
pri-miR-26a-2 REV	AAGAAGCGCACACACCATGT
pri-miR-34a FWD	CTGGGGAGAGGCAGGACA
pri-miR-34a REV	GGCATCTCTCGCTTCATCTT
pri-miR-126 FWD	AGCGCAGCATTCTGGAAGAC
pri-miR-126 REV	AGGGAGGTCAAGGCTGAGGT
pri-miR-199b FWD	GGGGTTCTCGGATCTCCA
pri-miR-199b REV	GTCGGTCCAGCTCTCCAGT
pri-miR-638 FWD	AGCTCGGTGCTGGAGAACTT
pri-miR-638 REV	GAAAGCTGGGACCAGATGGA
pri-miR-1228 FWD	GGGGCAGGTGTGTGGTG
pri-miR-1228 REV	GGCGAGGCAGGTGTGAG
pri-miR-3188 FWD	CCTTCTGGAGAGGCTTTGTG
pri-miR-3188 REV	GAGGGAGGGAAGGAAGGAG
pri-miR-17 FWD	AGAGAACATCACCTTGTA AAACTGAA
pri-miR-17 REV	CCTGCACTTTAAAGCCCAACT
pri-miR-18a FWD	CCTGCTGATGTTGAGTGCTTT
pri-miR-18a REV	ACACCTATATACTTGCTTGGCTTG
pri-miR-29b FWD	CCTTTCTAGGTTGTCTTGGGTTT
pri-miR-29b REV	TACTTCAGAGCTGTCCCATTCA
pri-miR-155 FWD	TTTATGCCTCATCCTCTGAGTG
pri-miR-155 REV	CAATTTGTTCCATGTGAATGCT
pri-miR-423 FWD	CGAAGTTTGAGGGAGAACTTG
pri-miR-423 REV	TCTAAGGAAGCCAGGAACTGTC
U3 FWD	CGAAAACACGAGGAAGAGA
U3 REV	CACTCCCCAATACGGAGAGA
U6 FWD	ATCTTGTTGGAAGGACGAAACA
U6 REV	AAATATGGAACGCTTCACGAAT
mt tRNAval FWD	CAGAGTGTAGCTTAACACAAAGCA
mt tRNAval REV	CAAGTTAAGTTGAAATCTCCTAAGTG
Exportin-5 FWD	AGTGGAAATCCAATCTTCCGTA
Exportin-5 REV	ATTTTGGCTAGCATTTCTGGTG
Exportin-1 FWD	AAAATACGTTGTTGGCCTCATT
Exportin-1 REV	GGGCCATTCTTGTTTCAGTATC
TGS1 FWD	GGATTCAAGTATGGCTCAGGAC
TGS1 REV	TGTCTGCGCATGTCTAGGTACT
p27 FWD	GGGGCTCGTCTTTTCGGGGTGTTT
p27 REV	GAGCGGGAGGGCGGAGAGGAG
GAPDH FWD	CCACTCCTCCACCTTTGAC
GAPDH REV	ACCCTGTTGCTGTAGCCA
c. elegans ama-1 FWD	GGAGCTCGAGTGGATCTTCG
c. elegans ama-1 REV	GCGCAGAGAGTATCCTGGAC
Drosha (exon 6) FWD	ACAACCGAAGATCACCATCTCT
Drosha (exon 6) REV	CATTCTTGTCAGGCATGGTCT
Drosha FWD	GAATGCAAAGGCATGATTGTTA
Drosha REV	AAAGTGGACGATAATCGGAAAA

Table S3. Primers used in this study for qRT-PCR.

Table S4

	Proliferating + siRNA CTRL	Proliferating + siRNA CTRL	Proliferating + siRNA TGS1	Proliferating + siRNA TGS1	Quiescent + siRNA CTRL	Quiescent + siRNA CTRL	Quiescent + siRNA TGS1	Quiescent + siRNA TGS1
	IP CTRL	IP TMG	IP CTRL	IP TMG	IP CTRL	IP TMG	IP CTRL	IP TMG
pri-miR-34a	2.35	39.84	6.22	15.56	1.62	147.25	0.09	35.08
pri-miR-3188	11.87	62.32	1.84	29.57	3.46	235.98	1.48	53.29
pri-miR-423	2.86	2.97	1.00	4.02	0.58	2.68	0.38	4.09

Table S4. Quiescence-induced pri-miRNAs contain a TMG cap. Amounts of pri-miR-34a, pri-miR-3188 and pri-miR-423 immunoprecipitated with anti-rabbit serum (IP CTRL) or antibody recognizing TMG-cap RNA (IP TMG) in proliferating or quiescent HFFs transfected with control siRNA or siRNA targeting TSG1. Relative expression was obtained by comparing input RNA and normalizing against a spike-in of non-endogenous RNA using qRT-PCR. Values in the table were multiplied by 1000 for clarity.

XII. Supplementary Figures

Figure S1

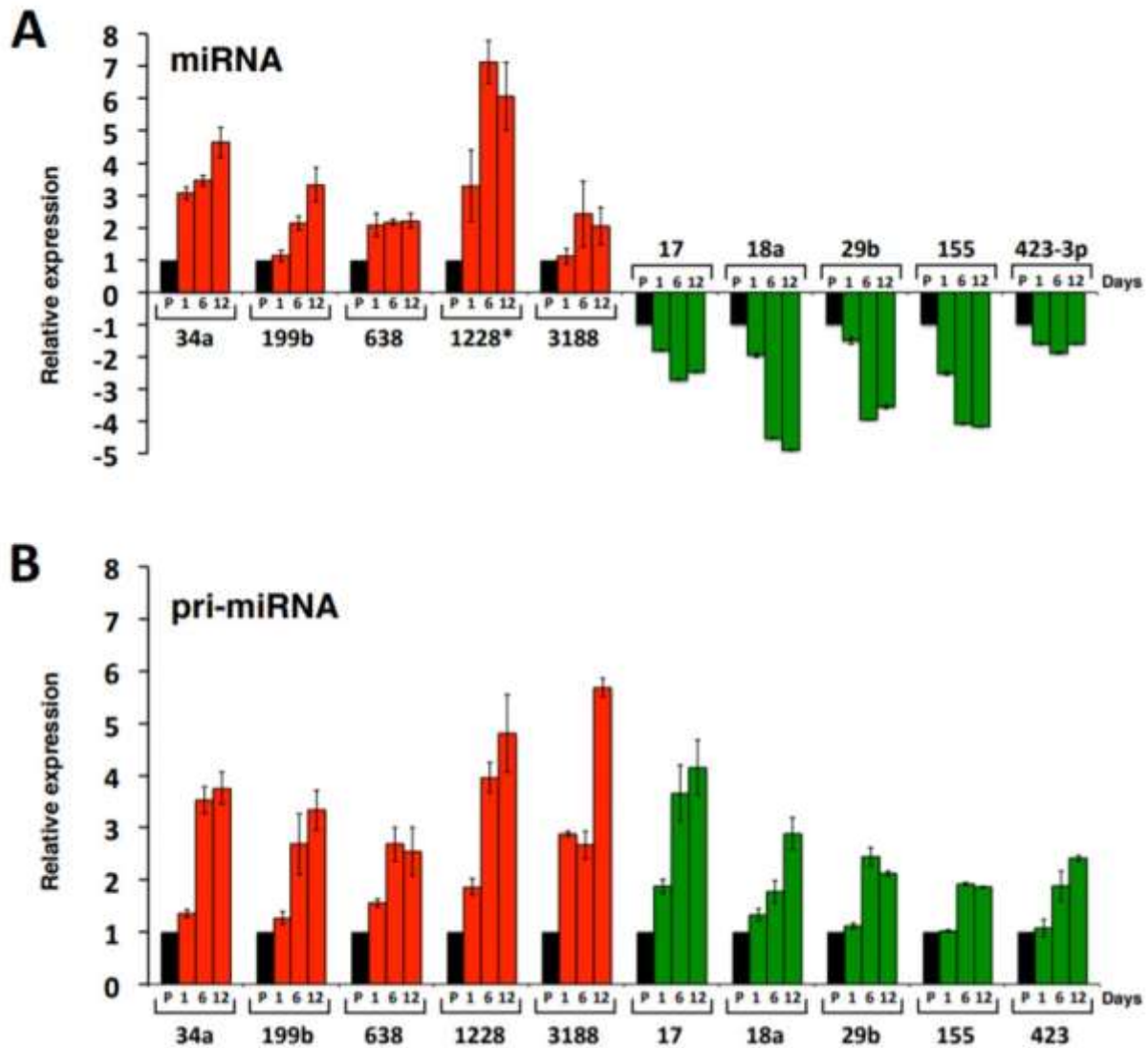


Figure S1. Differential expression of mature miRNAs and their corresponding primary miRNAs during the induction of quiescence by cell confluency. **(A)** qRT-PCR analysis of mature miRNAs. RNA was extracted 1, 6 and 12 days after reaching complete confluency and compared to RNA isolated from proliferating HFFs (P). **(B)** qRT-PCR analysis of primary miRNAs by using primers that bind upstream and downstream of the miRNA stem loop. In both A and B, red bars represent the miRNAs up-regulated during serum starvation by miRNA microarray analysis. Green bars represent mature miRNAs down-regulated during serum starvation by miRNA microarray analysis. Black bars represent miRNAs in proliferating (P) cells. Results were normalized to RNU43 and GAPDH mRNA expression and expressed relative to proliferating cells. Similar results were obtained in two independent experiments.

Figure S2

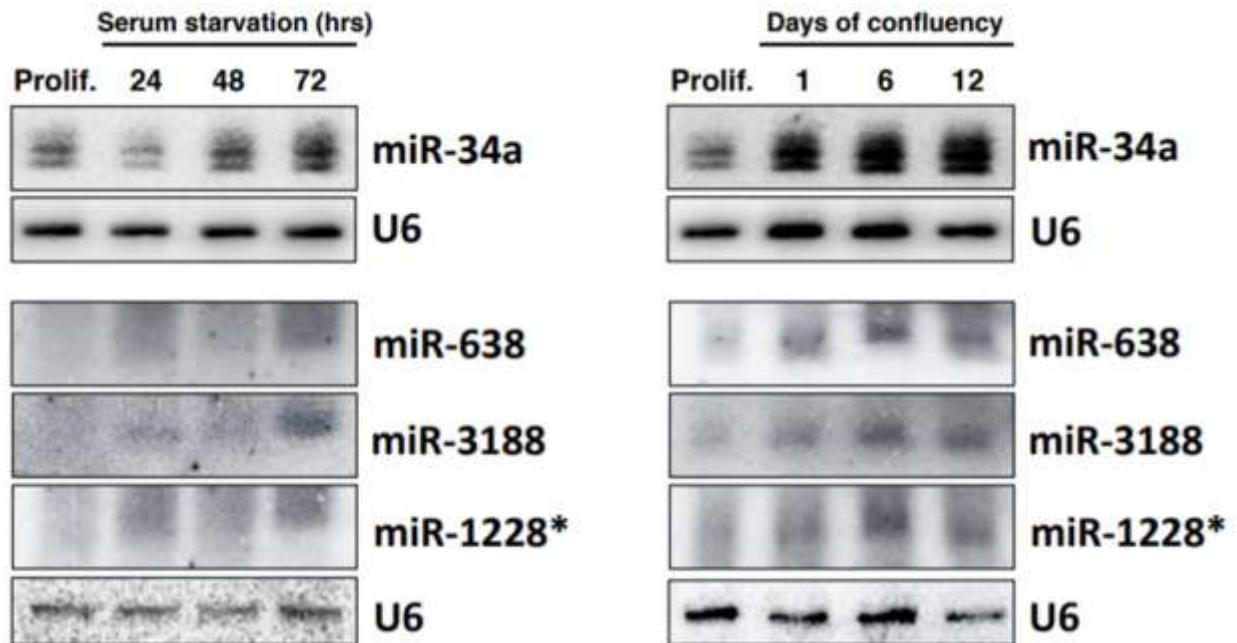


Figure S2. Northern blot analysis validating several miRNAs up-regulated during the induction of quiescence by serum starvation or cell confluency. U6 was used as loading control.

Figure S3

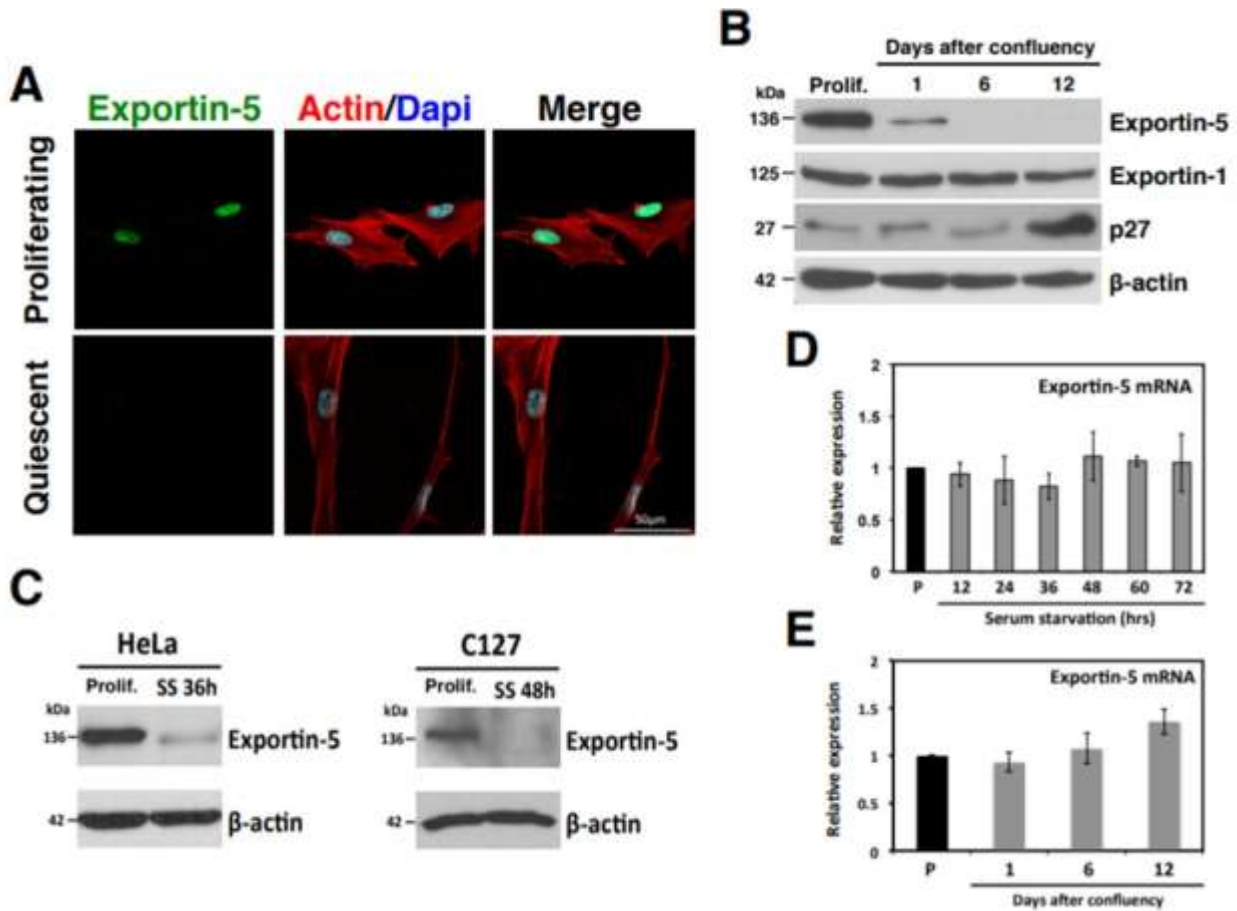


Figure S3. Exportin-5 expression is reduced post-transcriptionally during quiescence. **(A)** Immunofluorescence of Exportin-5 (green) in proliferating and quiescent HFFs by serum starvation. Nuclear fluorescence (DAPI, blue) and cytoskeleton immunofluorescence (actin, red) were used to visualize the nuclear localization of Exportin-5. **(B)** Western blot analysis showing reduction of Exportin-5 but not Exportin-1 levels after induction of quiescence by confluency. **(C)** Exportin-5 protein expression is markedly reduced during quiescence by serum starvation in cervical carcinoma (HeLa) and murine mammary tumor (C127) cell lines. β -actin was used as a loading control. **(D and E)** qRT-PCR analysis of Exportin-5 mRNA, showing no changes in levels during serum starvation or confluency compared to proliferating cells (P). Exportin-5 mRNA levels were normalized to GAPDH mRNA. Results show the average of three independent experiments.

Figure S4

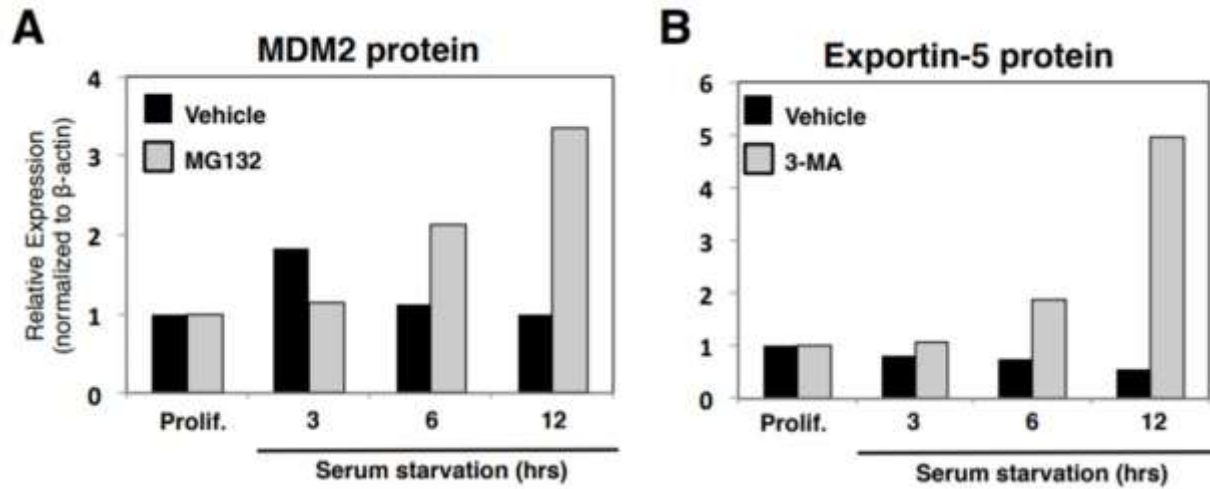


Figure S4. Accumulation of MDM2 and Exportin-5 caused by the inhibition of the proteasome or the autophagy pathways, respectively, during the induction of quiescence. **(A)** Western blot quantification of MDM2 protein in cells treated with vehicle (DMSO) or proteasome inhibitor MG132. **(B)** Exportin-5 protein in cells treated with vehicle (water) or autophagy inhibitor 3MA was measured by western blotting and quantified by using Image J software 1.48 (<http://imagej.nih.gov/ij/download.html>). Results were normalized to β -actin levels and compared to proliferating HFFs. Experiment was performed in triplicate with similar results. The figure shows a representative experiment.

Figure S5

	80	90	100
H. sapiens (Human)	CUUCUUG-	CACUGCCACCUCACUUUCCA	
P. troglodytes (Chimpanzee)	CUUCUUG-	CACUGCCACCUCACUUUCCA	
M. mulatta (Rhesus)	CUUCUUG-	CACUGCCACCUCACUUUCCA	
O. garnetti (Bushbaby)	CUACUCA-	CACUGCCACCUCACUUUCCA	
T. belangeri (Treeshrew)	CCUCUCA-	CACUGCCACUUCACUUUCCA	
M. musculus (Mouse)	CUCCUCG-	CACUGCCGCCUCACUUUCCA	
R. norvegicus (Rat)	UUCCUCA-	CACUGCCACCUCACUUUCCA	
C. porcellus (Guinea pig)	CCUUUUG-	CACUGCCACCUCACUGUCCA	
O. cuniculus (Rabbit)	CUUCUUA-	CACUGCCACCUCACCUCUCA	
S. araneus (Shrew)	CCUCUC--	CUCUGACACCUCACUUGCCA	
E. europaeus (Hedgehog)	CCUCUC--	CACUGCCACUUCACUUUGCA	
C. familiaris (Dog)	CCUCUC--	CACUGCCACCUCACUUUGUA	
F. catus (Cat)	CCUCUC--	CACCGCCACUUCACUUUCCA	
E. caballus (Horse)	CCUCUC--	CACUGCCGCCUCACUUUACA	
B. Taurus (Cow)	CCUCUC--	CACUGCCACCUCACUUUCCA	
		miR-34a BS	

Figure S5. The Exportin-5 3'UTR contains a conserved miR-34 binding site. Comparison of the Exportin-5 3'UTR (nucleotide 74 to 100) in different organisms, showing a conserved binding site for the miR-34 family (nucleotides in black box). Analysis was obtained using TargetScanHuman 6.2 software (www.targetscan.org).

Figure S6

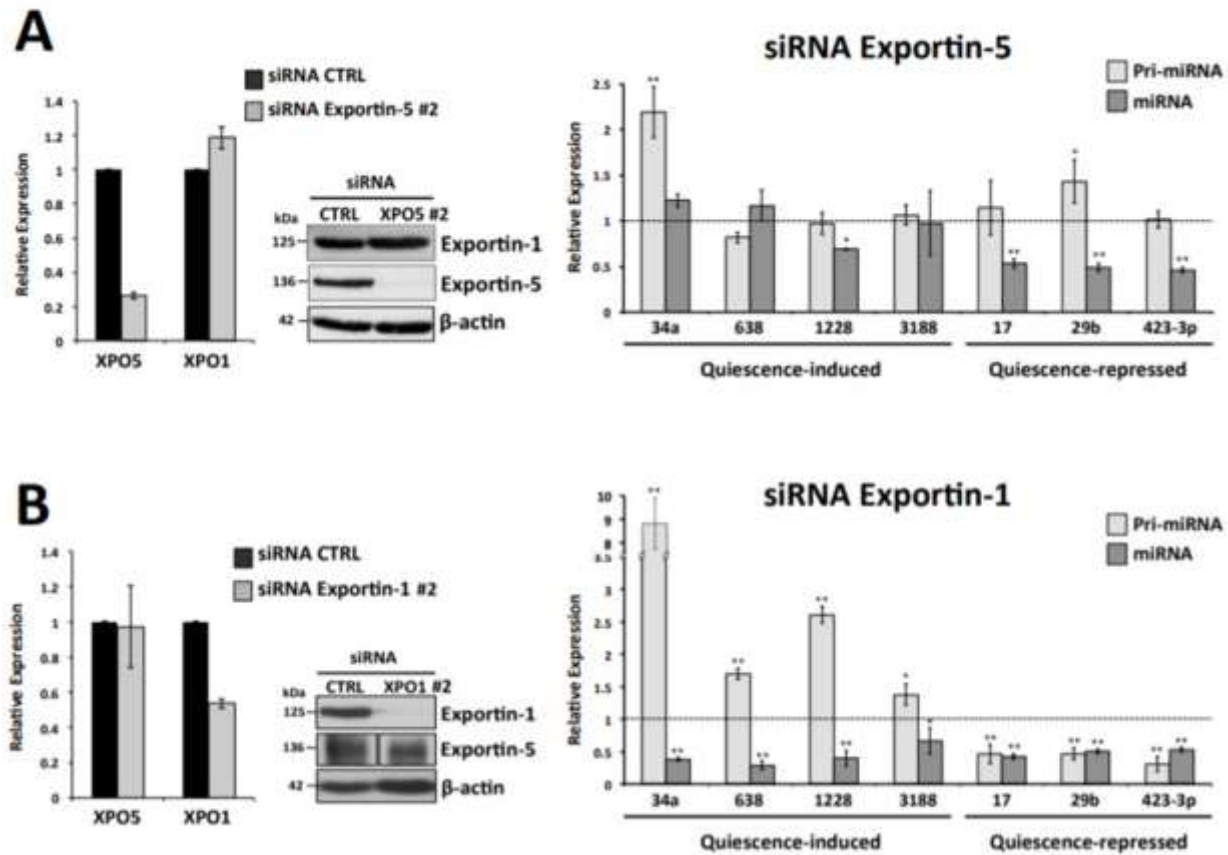
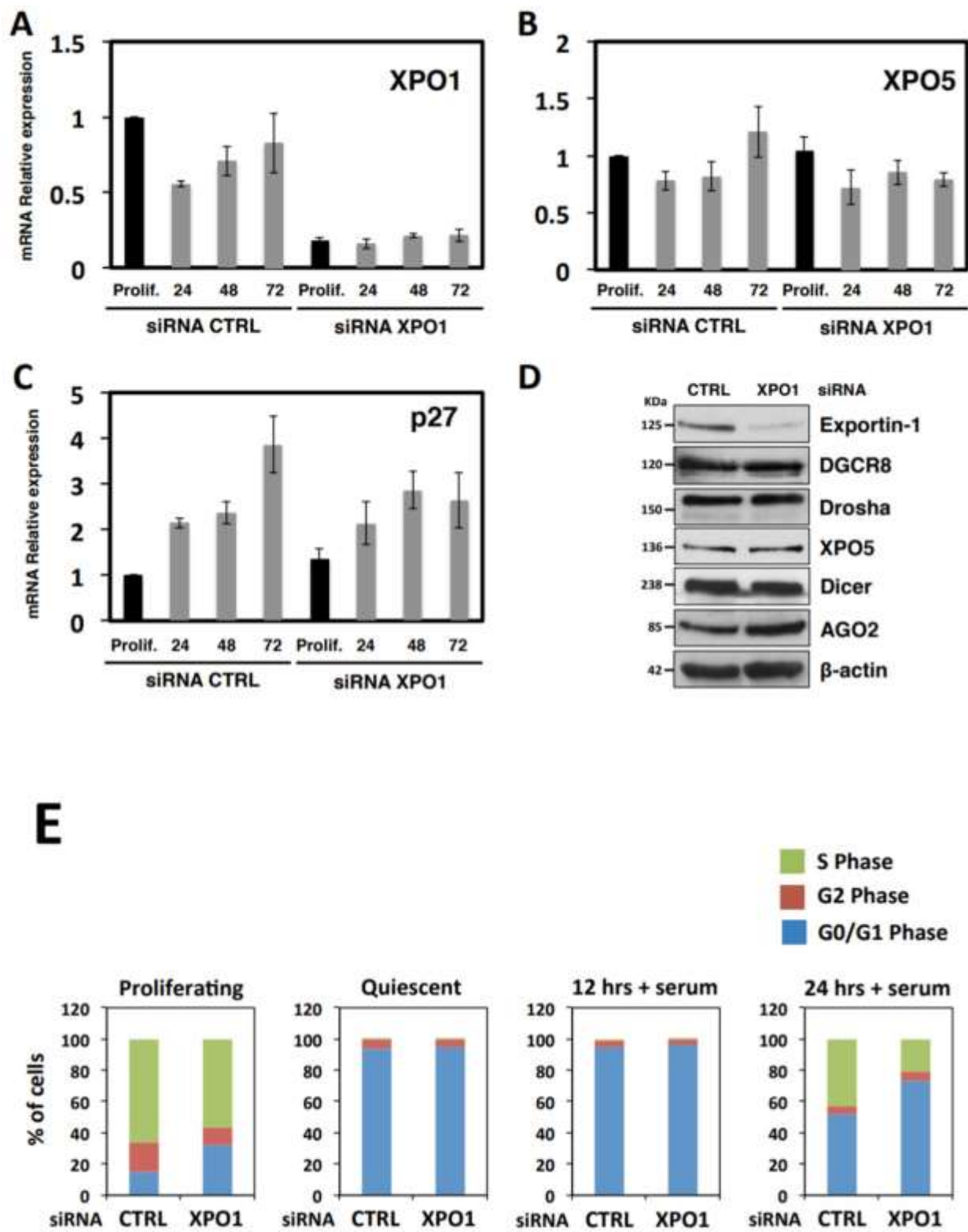


Figure S6. Biogenesis of quiescence-induced miRNAs requires Exportin-1 but not Exportin-5. **(A)** qRT-PCR analysis showing relative expression of quiescence-induced and quiescence-repressed primary miRNAs (light gray bars) and their corresponding mature miRNAs (dark gray bars) in proliferating HFFs 48 hrs after transfection with an siRNA scrambled control (dashed line) or an siRNA targeting Exportin-5 (siRNA XPO5 #2) (IDs33191, Ambion) (dark and light gray bars). Western blot and qRT-PCR analysis of Exportin-1 and Exportin-5 using XPO5-depleted or control HFF lysates confirms specific knock-down of Exportin-5. **(B)** qRT-PCR analysis showing relative levels of the RNAs tested in panel (A) in proliferating HFFs 48 hrs after transfection with an siRNA scrambled control (dashed line) or an siRNA targeting Exportin-1 (siRNA XPO1 #2) (ID118206, Ambion) (dark and light gray bars). Western blot and qRT-PCR analysis of Exportin-1 and Exportin-5 using XPO1-depleted or control HFF lysates confirm specific knock-down of Exportin-1. Two-tailed t-test results are indicated by * for $P < 0.05$ and ** for $P < 0.01$.

Figure S7



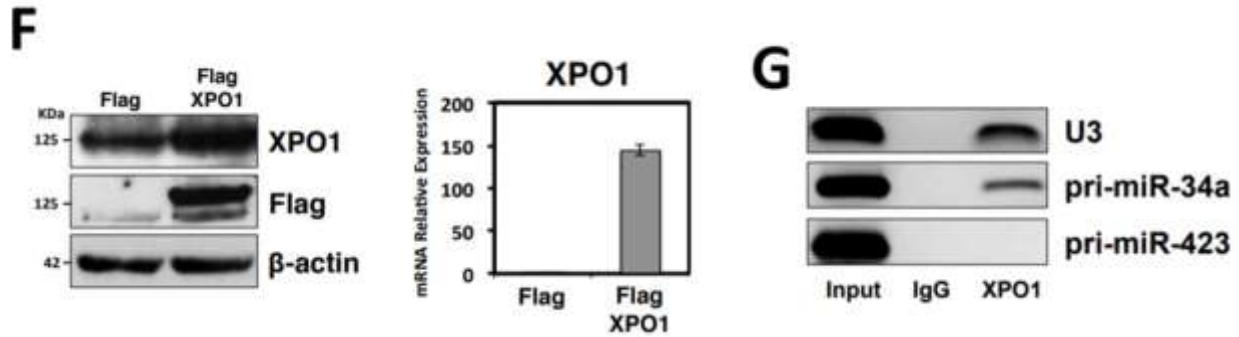


Figure S7. Exportin-1 knock-down during serum starvation does not alter the expression of Exportin-5, p27 or key proteins involved in miRNA biogenesis. HFFs were transfected with a siRNA scrambled control (CTRL) or a siRNA targeting Exportin-1 (XPO1) (ID s14937, Ambion) 24 hrs before the induction of quiescence by serum starvation. Expression of Exportin-1 (A), Exportin-5 (XPO5) (B), or p27 (C) mRNA in proliferating cells (Prolif) or at the indicated hours after serum starvation was determined by qRT-PCR and normalized to GAPDH mRNA. (D) Western blot showing repression of Exportin-1 but not other proteins involved in miRNA biogenesis in proliferating HFF cells transfected with an siRNA against Exportin-1 (XPO1). Similar results were obtained in two independent experiments. (E) Exportin-1 knock-down do not change the entrance of HFFs to quiescence. Cell cycle assays in proliferating, quiescent (24 hours after serum starvation), and synchronized (12 and 24 hours after the addition of serum) HFFs with siRNA scrambled control (CTRL) or a siRNA targeting Exportin-1 (XPO1). Similar results were obtained in two independent experiments. (F) Overexpression of Flag-Exportin-1. Western blot and qRT-PCR of HFF cells expressing Flag-tagged empty vector or Flag-tagged Exportin-1 (G) Exportin-1 physically associates with pri-miR-34a in quiescent cells. RT-PCR was used to detect U3 snoRNA, pri-miR-34a, or pri-miR-423 in control Flag-Vector (IgG) or Flag-Exportin-1 (XPO1) immunoprecipitates from quiescent HFFs. Similar results were obtained in two independent experiments.

Figure S8

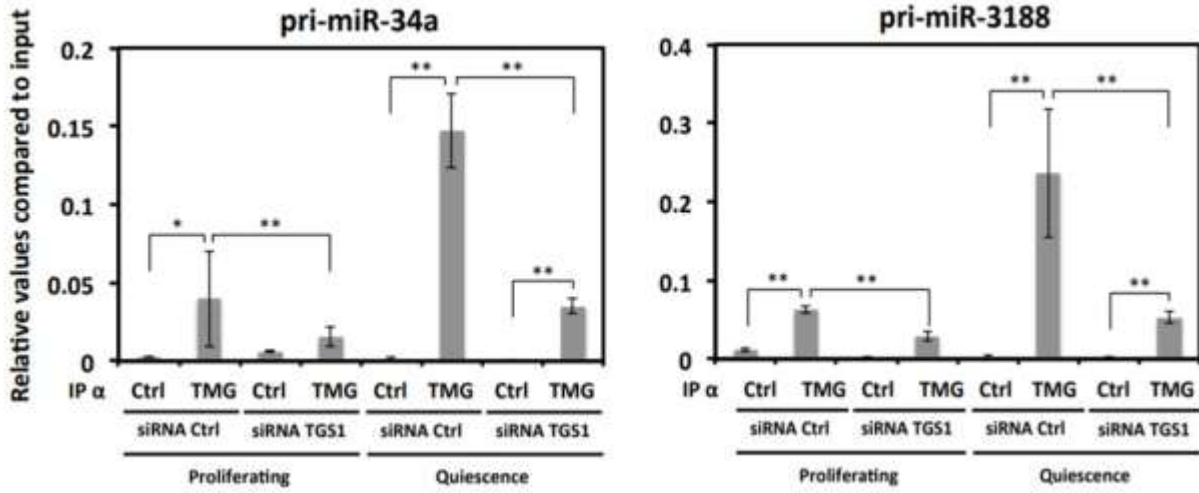


Figure S8. Quiescence-induced primary miRNAs are trimethylguanosine (TMG)-capped by TGS1. qRT-PCR of anti-TMG pull-downs as in Figure 4A to quantify levels of miR-34a and miR-3188. Relative expression was compared to input RNA and normalized against a spike-in of non-endogenous RNA. Two-tailed t-test results are indicated by * for $P < 0.05$ and ** for $P < 0.01$.

Figure S9

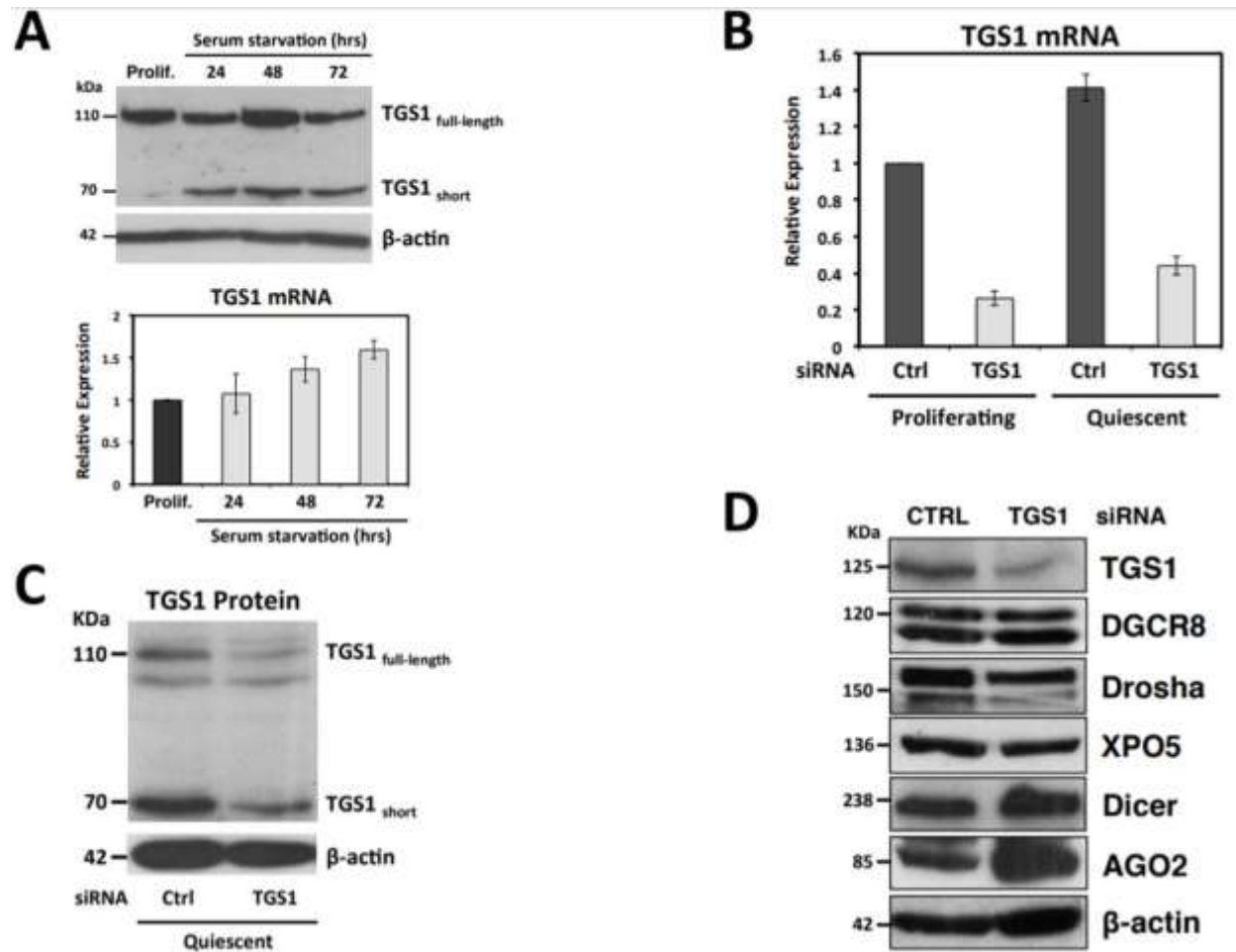


Figure S9. Analysis of TGS1 protein knock-down in proliferating and quiescent HFFs. **(A)** TGS1 expression increases during quiescence. Western blot (top panel) showing increased expression of the short isoform of TGS1 protein and qRT-PCR analysis (bottom panel) showing increased expression of total TGS1 mRNA in HFFs after serum starvation. **(B)** qRT-PCR analysis showing relative expression of TGS1 mRNA in proliferating and quiescent HFFs transfected with an siRNA scrambled control (dark grey bars) or an siRNA targeting TGS1 mRNA (Ambion) (light grey bars). **(C)** Western blot showing partial knock-down of TGS1 by using a siRNA against TGS1 (Ambion) in quiescent HFFs. **(D)** Western blot showing repression of full-length TGS1 but not other proteins involved in miRNA biogenesis in proliferating HFF cells transfected with a siRNA pool against TGS1 (Santa Cruz). Similar results were obtained in two independent experiments.

Figure S10

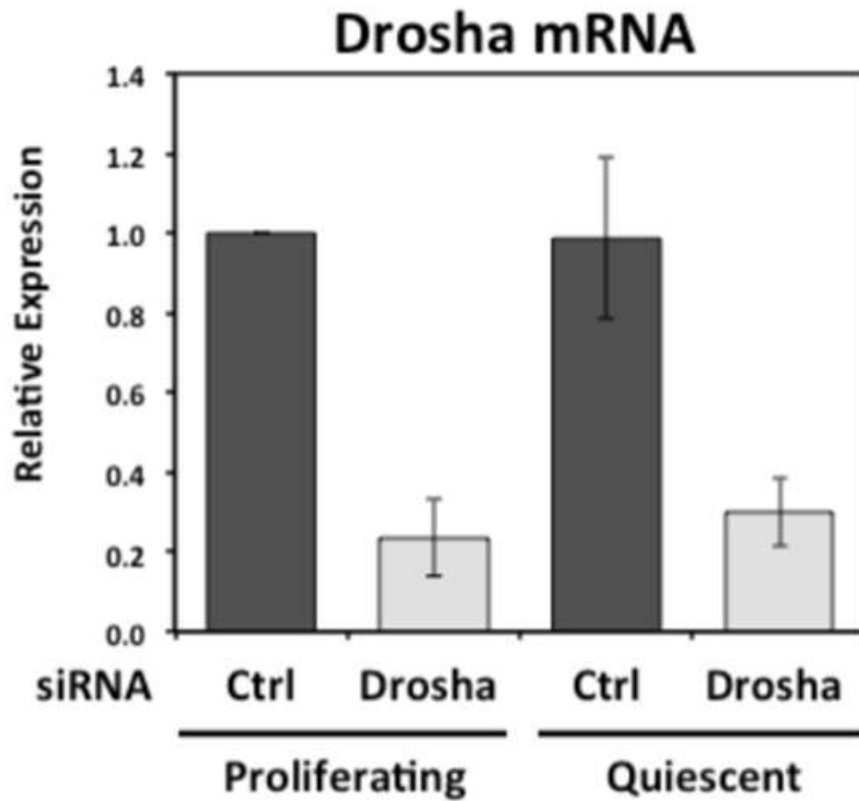


Figure S10. Analysis of Drosha knock-down in proliferating and quiescent HFFs. qRT-PCR analysis showing relative expression of Drosha mRNA in proliferating and quiescent HFFs transfected with siRNA scrambled control (dark grey bars) or siRNA pool targeting Drosha (Santa Cruz) (light grey bars). Cells were made quiescent by serum starvation for 48 hrs. Similar results were obtained in three independent experiments. The figure shows a representative experiment.

Figure S11

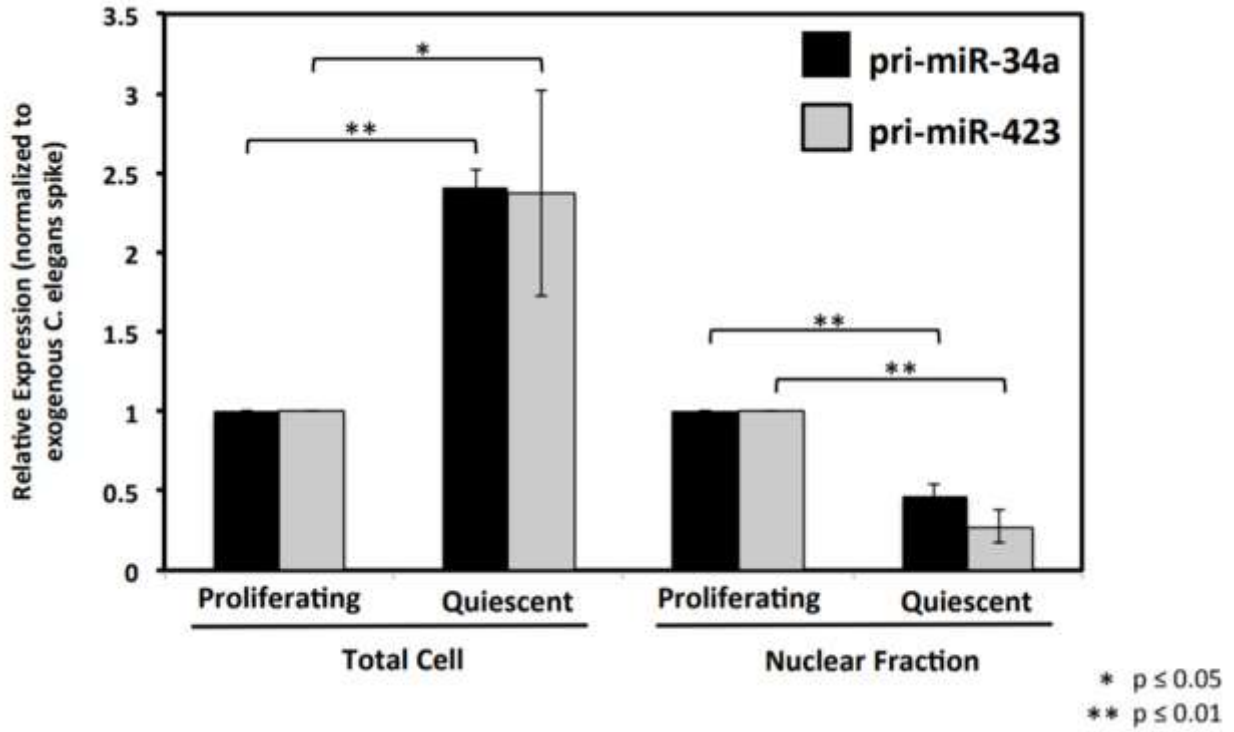


Figure S11. Comparison of pri-miRNAs present in total cell versus nuclear fraction. qRT-PCR analysis shows a significant reduction of pri-miR-34a and pri-miR-423 in nuclear fractions in comparison to total cell in quiescent HFF cells. Results show the average of three independent experiments. Two-tailed t-test results are indicated by * for $P < 0.05$ and ** for $P < 0.01$.

Figure S12

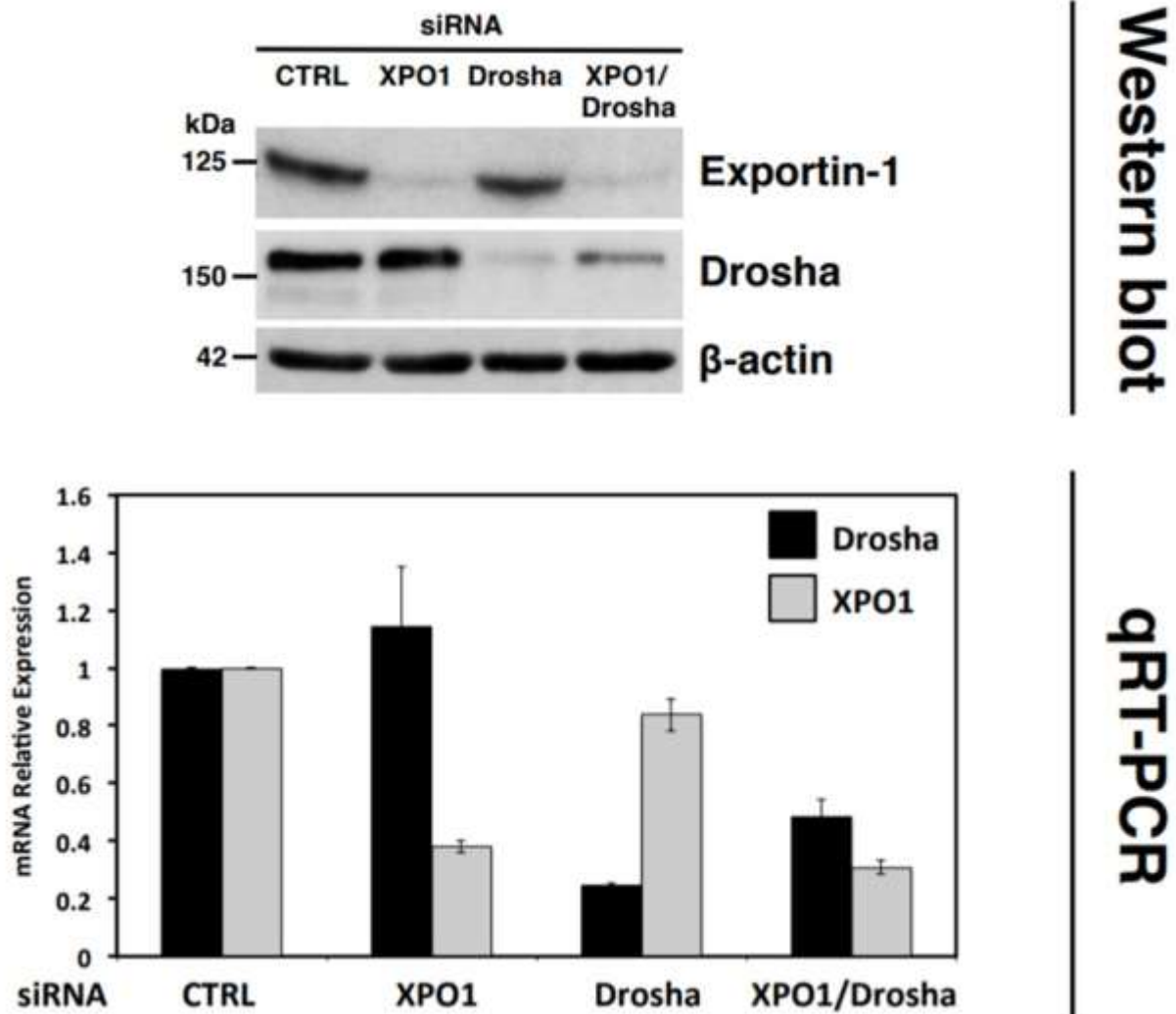


Figure S12. Verification of double knock-down of Drosha and XPO1. Knock-down of Drosha (Santa Cruz) and/or Exportin-1 (ID s14937, Ambion) was performed by using siRNAs against these proteins in proliferating HFFs and confirmed by Western blot analysis and qRT-PCR. siRNA control was added to XPO1 only and Drosha only siRNA transfections to equalize total siRNA amounts. Similar results were obtained in two independent experiments.

Figure S13

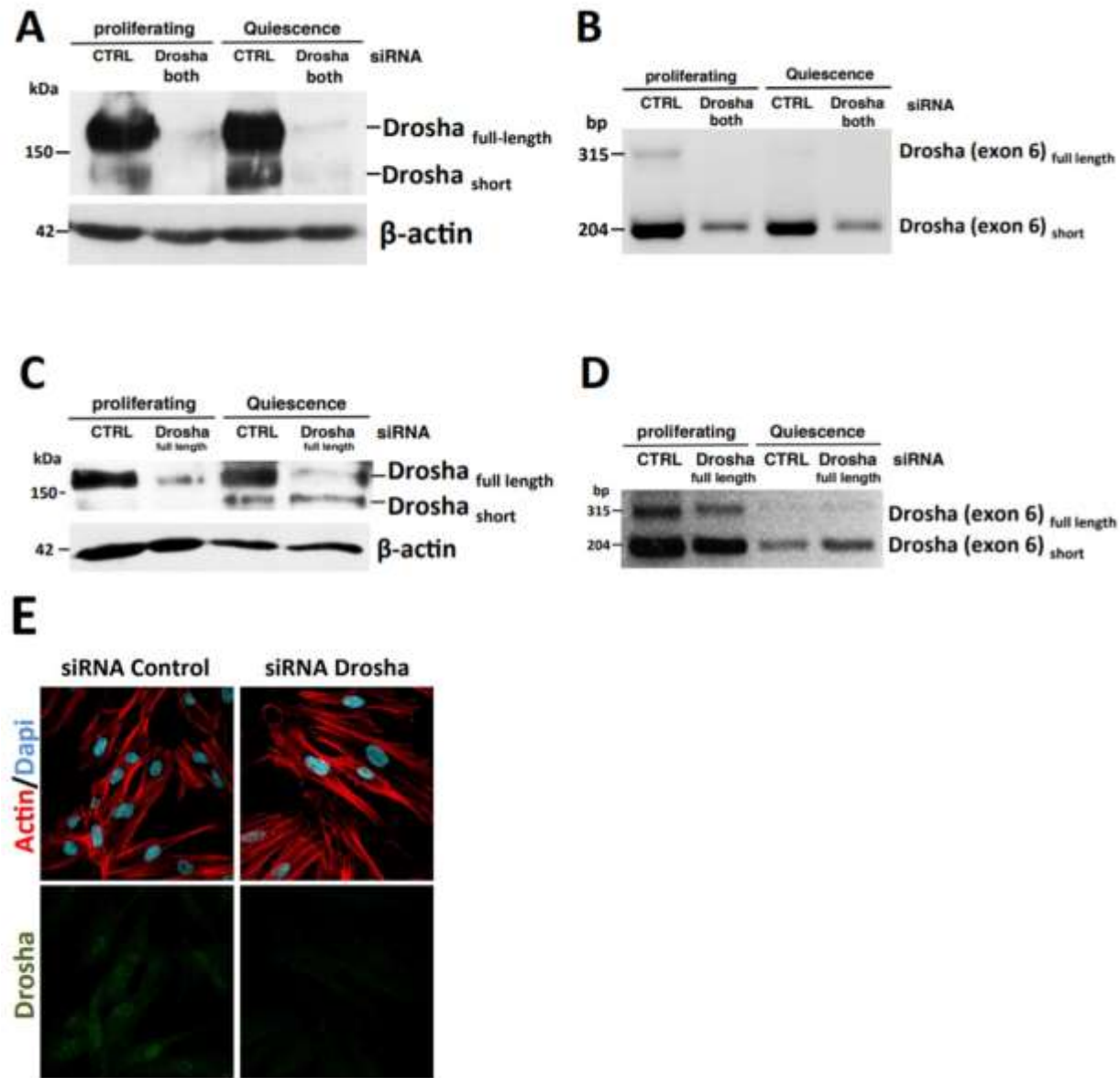


Figure S13. Transient knock-down of Drosha isoforms. **(A)** Western blot analysis of Drosha in proliferating and quiescent HFF cells 72 hrs after knock-down of both Drosha isoforms by using a pool of siRNAs against different regions of Drosha mRNA (Drosha pool, Santa Cruz). **(B)** qRT-PCR of the exon 6 region of Drosha mRNA (204bp for the short Drosha and 315bp for the full length Drosha) after knock-down of both isoforms of Drosha. **(C)** Western blot analysis of Drosha in proliferating and quiescent HFF cells 72 hrs after selective knock-down of full length Drosha isoform only by using a siRNAs targeting exon 4 of Drosha mRNA (Drosha full length) that is absent (alternative splicing) from the short isoform of Drosha. **(D)** qRT-PCR of the exon 6 region of Drosha, as detailed in (B), in proliferating and quiescent HFF cells 72 hrs after specific knock-down of Drosha full length. **(E)** To verify specificity of Drosha antibody for immunofluorescence, confocal microscopy images were obtained using the same settings. HFF treated with siRNA Drosha (Drosha pool, Santa Cruz) demonstrate a loss of Drosha staining (pseudo-color green). Nuclear (DAPI, blue) and cytoskeleton (actin, red) immunofluorescent staining were used to visualize the cellular localization of Drosha.

Figure S14

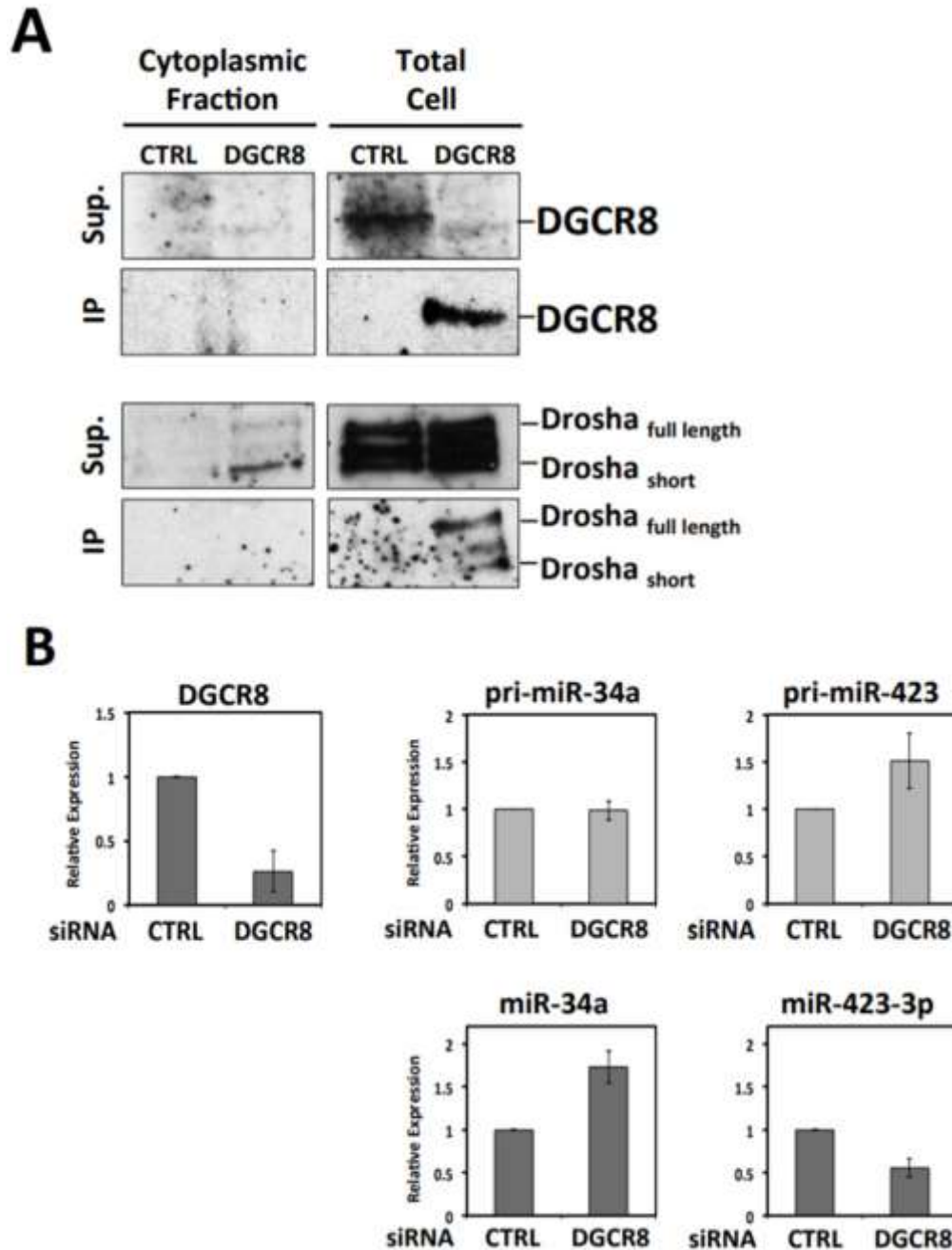


Figure S14. DGCR8 is not present in the cytoplasm of quiescent HFFs and the processing of quiescent-induced miR-34a is not affected by DGCR8 knock-down (A) DGCR8 is not interacting with the cytoplasmic isoform of Drosha. Immunoprecipitation (IP) of DGCR8 in quiescent HFFs shows absence of this protein in the cytoplasmic fraction and no pull-down of the small isoform of Drosha. DGCR8 polyclonal antibody (Proteintech) was used for the IP and rabbit IgG as a negative control (CTRL). Supernatant (Sup.) is also shown as an IP control. Similar results were obtained in two independent experiments. (B) miR-34a processing is not affected by DGCR8 knock-down in quiescent HFFs. Knock-down of DGCR8 was performed by using siRNAs against these proteins in quiescent HFFs and confirmed by qRT-PCR. siRNA scramble was used as a control (CTRL). Similar results were obtained in two independent experiments.

Figure S15

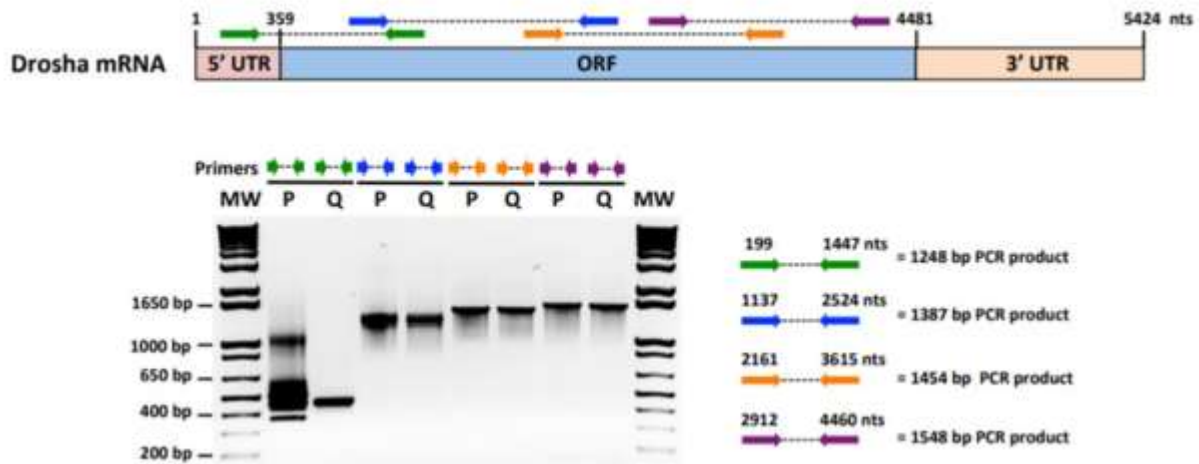


Figure S15. Detection of a short isoform of Drosha in quiescent HFF cells. Tiling RT-PCR of The Drosha mRNA open reading frame (ORF). Colored arrows represent different pairs of primers at different positions in Drosha mRNA with the predicted sizes of the wild-type amplification products shown. P indicates proliferating HFF cells, Q indicates quiescent HFF cells, and MW indicates molecular weight ladder. Similar results were obtained in three independent experiments.

Figure S16

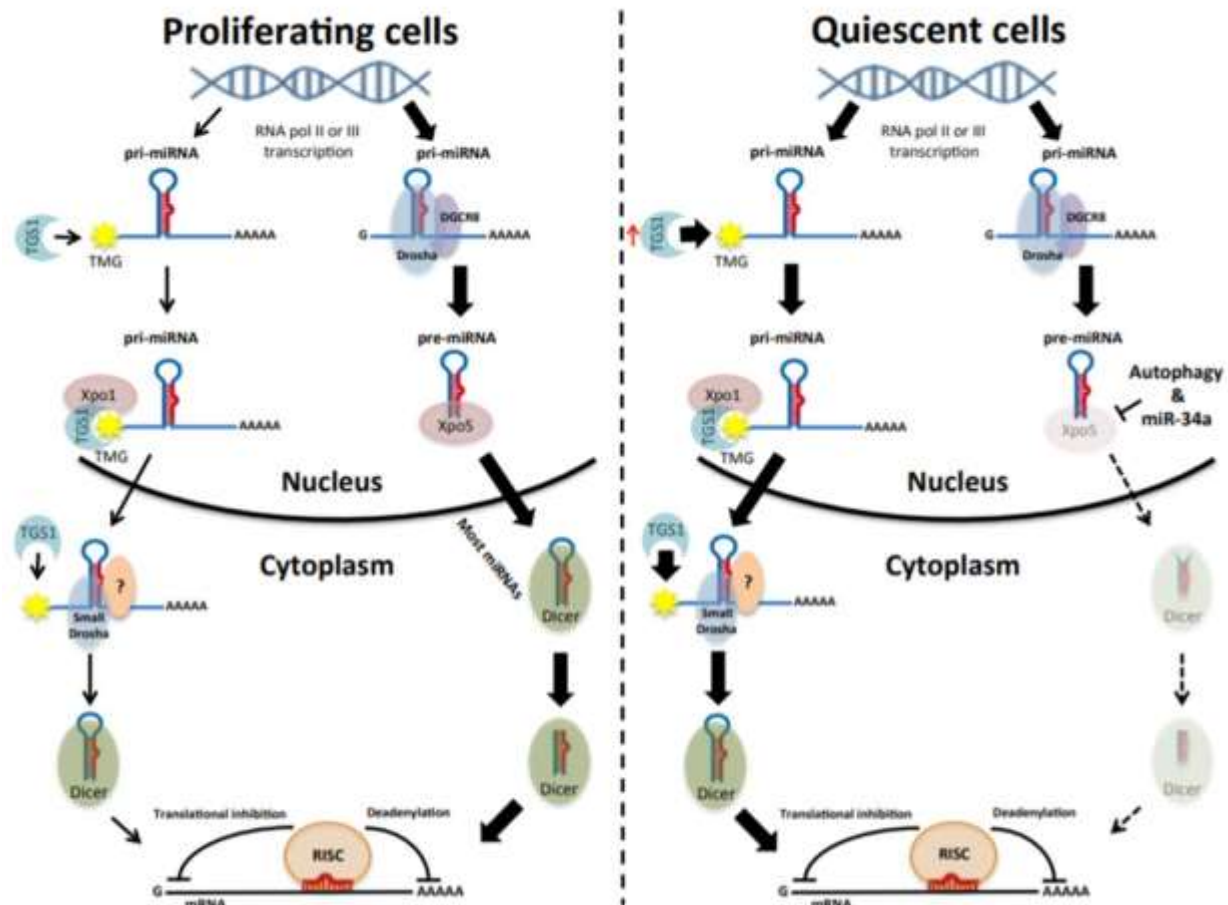


Figure S16. Model showing an alternative miRNA biogenesis pathway of (TMG)-capped pri-miRNAs. In proliferating cells, most miRNAs are processed through the canonical Exportin-5-dependent miRNA biogenesis pathway, but the (m⁷G)-cap of a small group of pri-miRNAs is hypermethylated by TGS1 to form a (2,2,7-TMG)-cap, allowing recognition by Exportin-1. In quiescent cells, markedly reduced expression of Exportin-5, due to autophagy and miR-34a-mediated repression, limits the processing of most miRNAs, but increased expression of TGS1 supports the production of (TMG)-capped pri-miRNAs, which can be exported by Exportin-1 to the cytoplasm. A short form of Drosha in the cytoplasm (possibly exported from the nucleus by Exportin-1) processes the (TMG)-capped pri-miRNAs.

XII. Acknowledgements

We thank Demian Cazalla, Brian Adams, Kasandra Riley, Antonio Giraldez, Anne Edwards, Sara Marlatt, and Jan Zulkeski at Yale University for advice and assistance. We also thank Dr. Arjumand Ghazi of the University of Pittsburgh, and Dr. Frank Slack of Harvard University, for their support with reagents. Microarray processing was carried out in the Yale Center for Genome Analysis (YCGA). Imaging experiments and image analysis were performed in the West Virginia University (WVU) Microscope Imaging Facility, supported by the WVU Cancer Institute and NIH grants P20 RR016440, P30 GM103488, and P20 GM103434. I.M., J.B., and K.H. were supported in part by a WVU Foundation Fund (2V882). K.H. was supported in part by funding from the Ladies Auxillary to the VFW of the United States (CK 003229). S.I.A.B. and S.V. were supported by Massachusetts General Hospital startup funds. M.X. is supported by NIH grant R00 CA190886. J.A.S. is an Investigator of the Howard Hughes Medical Institute. This work was supported by Program Project Grant CA016038 from the National Cancer Institute (D.D. and J.A.S.).

XIII. Supplementary Materials and Methods

A. Cell culture and reagents

Normal primary human foreskin fibroblasts (HFFs, obtained from the Yale Skin Disease Research Center), HeLa/E6-5K, 293T, and C127 cells were cultured in DMEM (Sigma-Aldrich, St. Louis, MO) supplemented with 10% fetal bovine serum (FBS), L-glutamine, 10 mM Hepes pH 7.3, penicillin–streptomycin, and amphotericin B (DMEM-10) at 37°C in the presence of 5% CO₂.

B. Quiescence induced by serum starvation and confluency

HFFs, HeLa, and C127 cells were rinsed with Dulbecco's phosphate buffered saline (DPBS, Life Technologies, Grand Island, NY), detached using minimal amounts of 0.25% trypsin-EDTA (Life Technologies) and plated in serum-free DMEM supplemented with MEM non-essential amino acids at low density (approx. 1×10^6 cells/150cm² flask). After various times at 37°C, cells were collected by scraping, pelleted by centrifugation, and rinsed twice with phosphate buffered saline (PBS, Sigma). For proliferating samples (0 h), HFFs, HeLa, and C127 cells were plated in DMEM supplemented with 10% FBS and MEM non-essential amino acids at a density of approximately 1.5×10^6 cells/100mm plate, cultured for 24-48 h, scraped, and pelleted. For quiescence induced by confluency, we plated HFFs (2×10^6 cells/100 mm plate)

in DMEM-10. After cells reached 100% confluency (day 0), DMEM-10 was replaced every second day. Cells were scraped and pelleted by centrifugation at days 1, 6, and 12. Proliferating HFFs (Mock, day 0) were plated at 1×10^6 cells/100 mm plate in DMEM-10 and harvested at approximately 80% confluency.

C. Immunoprecipitation of tri-methylated capped RNA

Proliferating HFFs (2×10^6 cells/100 mm plate) were transiently transfected with 500 pmol siRNA TGS1 (ID S14937, Ambion) or siRNA negative control #1 (Ambion) using Lipofectamine RNAiMAX (Invitrogen) and incubated at 37°C for 24 h in antibiotic-free DMEM-10. Cells were then trypsinized, plated and incubated for 72 h in serum-free, antibiotic-free DMEM supplemented with MEM non-essential amino acids to induce quiescence or in the same medium supplemented with 10 % FBS for 24 h to maintain proliferation. RNA was extracted from pelleted cells using Trizol Reagent (Life Technologies) per manufacturer's instructions and treated with Turbo DNase (Ambion) for 20 min at 37°C. RNA concentration was determined with a Nanodrop 2000 Spectrophotometer (Thermo Scientific, Wilmington, DE). 1 μ L of RNasin Plus RNase inhibitor (Promega, Madison, WI) was added to 10 μ g of RNA and incubated for 3-5 min at 85°C; an additional 1 μ L of RNase inhibitor was added after heating. RNA diluted in NET-2 buffer (150 mM NaCl; 50 mM Tris, pH 7.4; and 0.05% NP-40) was incubated for 2 h at 4°C with 40 μ L/10 μ g RNA of Protein G Sepharose 4 Fast Flow beads (GE Healthcare, Buckinghamshire, UK). The beads were then removed by centrifugation. Pre-cleared RNA was rotated overnight at 4°C with 40 μ L Protein G Sepharose 4 Fast Flow beads loaded with 15 μ L normal rabbit serum (control, EMD-Millipore, Billerica, MA) or 15 μ L rabbit m3G-cap antibody (Synaptic Systems, Goettingen, Germany). Beads were rinsed 5 times with NET-2 buffer and resuspended in G-50 buffer (20 mM Tris, pH 7.5; 300 mM NaOAc; 2 mM EDTA; and 0.25% SDS). RNA was extracted from the beads by phenol-chloroform-isoamyl alcohol extraction and resuspended in 20 μ L of nuclease-free water. 1 μ L of GlycoBlue Coprecipitant (Ambion) was added to aid in visualization of the RNA pellet. Immunoprecipitated tri-methylated capped RNA was analyzed by RT-PCR and qRT-PCR for pri-miRNA. Relative expression was normalized using Solaris RNA spike kit (Thermo Scientific, Rockford, IL).

D. Microarray analysis

For global miRNA expression, we used the GeneChip® miRNA 2.0 Array (Affymetrix, Santa Clara, CA, USA) that assesses 1,105 mature human miRNAs. Microarray processing was carried out in the Yale Center for Genome Analysis (YCGA). Briefly, 1 μ g of total RNA was labeled using the Flashtag RNA labeling kit (Genisphere, Hatfield, PA, USA) according to the manufacturer's instructions. Each sample

was hybridized to the array at 48°C and 60 rpm for 16 hours then washed and stained on Fluidics Station 450 (Fluidics script FS450_0003) and finally scanned on a GeneChip® Scanner 3000 7G (Affymetrix, Santa Clara, CA, USA). Feature intensities were extracted by using miRNA 2.0 array library files and analyzed with the miRNA QC Tool software. Array data were deposited into ArrayExpress: E-MTAB-2760.

E. Quantitative RT-PCR

Total RNA from HFFs was extracted using Trizol Reagent (Life Technologies) per manufacture instructions and treated with Turbo DNase (Ambion) for 20 minutes at 37°C. RNA concentrations were determined with a Nanodrop 2000 Spectrophotometer (Thermo Scientific, Wilmington, DE). For mature miRNAs, 1 µg of total RNA was reverse transcribed using the TaqMan® miRNA Reverse Transcription Kit (Life Technologies) according to the manufacturer's instructions. For qRT-PCR of individual mature microRNAs, we used specific TaqMan® miRNA Assays (Life Technologies) analyzed on a Real-Time thermocycler iQ5 (Bio-Rad, Hercules, CA) according to the manufacturer's protocol. To measure pre-miRNAs, pri-miRNAs and mRNA levels, total RNA was converted to cDNA using the iScript cDNA synthesis kit (Bio-Rad), followed by RT-PCR or qRT-PCR using the iQ SYBR green supermix (Bio-Rad). Relative expression was calculated using the double delta CT method (relative expression = $2^{-\Delta\Delta CT}$; where $\Delta CT = CT (\text{Target RNA}) - CT (\text{endogenous control RNA})$), where the endogenous control for miRNAs was the small nucleolar RNA RNU43 and for pre-miRNAs, pri-miRNAs and mRNA was GAPDH. For qRT-PCR normalization by exogenous spike, we used *C. elegans* total RNA for the amplification of the worm-specific ama-1 gene. Primer sequences are shown in supplementary Table 4.

F. Western blot analysis

Proliferating or quiescent HFFs were lysed in NP-40 lysis buffer (50 mM Tris-HCl pH 8.0, 150 mM NaCl, 1 mM EDTA, 1% NP-40, 0.1% SDS, 0.5 mM sodium metavanadate, Protease Inhibitor Cocktail (Sigma-Aldrich) for 12 minutes at 4°C and clarified by centrifugation at 4°C (5 min., 13,000 rpm). The concentration of extracted proteins was determined by using the Pierce BCA Protein Assay (Thermo Scientific, Rockford, IL). Cell lysates (15-30 µg of total cell protein) were separated by SDS-PAGE, transferred to Immobilon-Fl membranes (Millipore, Bedford, MA) and blocked with 5% milk-TBST buffer (5% nonfat dried milk, 25 mM Tris-HCl pH 8.0, 125 mM NaCl and 0.5% Tween-20) for 1 h at room temperature. Membranes were incubated overnight at 4°C with primary antibodies, washed with TBST and incubated with appropriate horseradish peroxidase (HPR)-conjugated secondary antibodies in 5% milk-

TBST buffer for 1 h at room temperature [1:10.000; Pierce goat anti-mouse or Pierce donkey anti-rabbit (Thermo Scientific), anti-Rat IgG HRP-linked antibody (Cell Signaling), or donkey anti-goat (Santa Cruz Biotechnologies, Dallas, TX)]. After rinsing with TBST, proteins were detected using the Pierce SuperSignal West Pico Chemiluminescent Substrate or Pierce SuperSignal West Femto Maximum Sensitivity Substrate (Thermo Scientific, Rockford, IL). The following primary antibodies were used: goat anti-Actin (C-11, 1:5000, Santa Cruz Biotechnologies), rabbit anti-MDM2 (1:250, Santa Cruz Biotechnologies), rabbit anti-p27 (C-19, 1:1000, Santa Cruz Biotechnologies), rabbit anti-p62/SQSTM1 (NBP1-49954, 1:1000, Novus Biologicals, Littleton, CO), rabbit anti-TGS1 (1:1000, Proteintech), mouse anti-Exportin-1 (C-1, 1:1000, Santa Cruz Biotechnologies), mouse anti-Exportin-5 (1:1000, Abnova, Walnut, CA), goat anti-Exportin-5 (E-16, 1:1000, Santa Cruz Biotechnologies), rabbit anti-Drosha (1:1000 ab12286, Abcam, Cambridge, MA), rabbit anti-Drosha (1:1000 D28B1, Cell Signaling Technology, MA), rabbit anti-DGCR8 (1:1000 10996-1-AP, Proteintech), rabbit anti- β -Tubulin (1:1000 9F3, Cell Signaling, Danvers, MA), rat anti-AGO2 (1 μ g/mL clone 11A9, Sigma-Aldrich), rabbit anti-Dicer (1:1000, Cell Signaling), and rabbit anti-HDAC1 (1:1000 H52, SantaCruz Biotechnologies).

G. Immunofluorescence and signal intensity quantitation

HFFs were plated on poly-L-lysine-coated glass coverslips at 5×10^4 cells/coverslip in serum-free DMEM supplemented with MEM non-essential amino acids (quiescent) and at 2.5×10^4 cells/coverslip in DMEM supplemented with 10% FBS and MEM non-essential amino acids (proliferating). Cells were rinsed with PBS and fixed with 10% buffered formalin phosphate (Fisher Scientific, Fair Lawn, NJ) for 15 minutes at room temperature 24 h, 48 h, and 72 h after plating; proliferating cells were fixed at 24 h. Cells were permeabilized for 10 minutes at room temperature with 0.4% Triton X-100 (Amresco, Solon, OH) in PBS and incubated with mouse anti-Exportin-5 (1:500) or rhodamine-conjugated phalloidin (F-actin, 1:500; Molecular Probes, Eugene, OR) in 5% bovine serum albumin (BSA) in PBS for 1 h at room temperature. Immunostained cells were then incubated for one h at room temperature with Alexa Fluor 488-conjugated goat anti-mouse secondary antibodies (1:2000; Molecular Probes) in 5% BSA in PBS. Cells were mounted on glass slides with ProLong Gold Anti-fade Reagent with Dapi (Life Technologies, Eugene, OR) to visualize nuclei. Zeiss LSM510 confocal microscope using AIM software (Carl Zeiss MicroImaging, Thornwood, NY, USA) was utilized to acquire fluorescent images. For immunofluorescence quantification, HFF cells were plated and cultured on poly-L-lysine coated coverslips (Neuvitro Corp., El Monte, CA) as described above. To diminish masking of cytoplasmic signal with nuclear staining, HFF cells were incubated in 3 μ g/mL of Digitonin (Promega, Madison, WI) in Buffer B (20 mM Hepes, 110 mM Potassium Acetate, 5 mM Sodium Acetate, 2 mM magnesium Acetate, 1 mM

EGTA and 2 mM DTT) for 3 min to selectively permeabilize only the plasma membrane (<http://www.nature.com/protocolexchange/protocols/1994#/figures>). After fixation with 3.7% formaldehyde in PBS, HFF cells were blocked with 5% BSA in PBS for 1 h at room temperature and immunolabeled as described above with rabbit anti-Drosha (D28131; 1:500; Cell Signaling) and Rhodamine-Phalloidin and mounted on glass slides with ProLong Gold Anti-fade Reagent with Dapi (Life Technologies). Confocal microscopy was used to visualize cells; fluorescent images were converted to TIF Files and localization of Drosha in the cytoplasm was quantitated utilizing Image J (<http://www.macbiophotonics.ca/imagej/>). In short, the area (outline) of the entire cell and nucleus were determined using F-actin and Dapi images, respectively, by the “Analyze/Analyze Particle” command after the “Threshold” was set for the images. “Nuclear” Drosha and “Total Cell” Drosha were measured by total pixel area from threshold Drosha image contained in the “outline” of the cell and nucleus (identified regions of interests, ROIs, obtained from F-actin and Dapi images). Quantification of “cytoplasmic” Drosha was generated by the subtraction of “Nuclear” Drosha due to partial permeabilization of the nuclear fraction with digitonin from “Total Cell” Drosha. Results were normalized to proliferating HFF cells; experiments were performed in triplicate.

H. Inhibition of proteasome activity and autophagy

HFFs plated the day before at 1.5×10^6 cells/100 mm plate were pre-treated with 10 μ M MG132 (2-Leu-Leu-H Aldehyde, Peptide International, Louisville, KY) or vehicle (DMSO) for 2 h at 37°C. Treated cells were trypsinized and plated at 1×10^6 cells/T150 flask in serum-free DMEM supplemented with MEM non-essential amino acids containing 10 μ M MG132 to induce quiescence. Cells were harvested by scraping at 3, 6, and 12 h and pelleted. Alternatively, MG132-treated HFF cells were scraped and pelleted after pretreatment for 0 h time point (mock). Pelleted cells were lysed and analyzed by Western blotting. For the 3-MA treatment, HFFs were plated at 1.5×10^6 cells/100 mm plate in DMEM-10 and incubated overnight. Cells pre-treated with 10 mM 3-methyladenine, 3-MA (Sigma-Aldrich) or vehicle (water) for 2 h were induced to quiescence by trypsinizing and plating at 0.9×10^6 cells/T150 flask and serum starved in the presence of 10 mM 3MA. Cells were scraped at 3, 6, and 12 h and pelleted to collect. Additionally, Mock (0 h) was generated by pelleting 3MA-treated proliferating HFFs. Harvested cells were lysed with NP40 lysis buffer and analyzed by Western blotting.

I. Transient transfection of siRNAs and mimicRNAs

75 μ L of Lipofectamine RNAiMAX (Invitrogen, Grand Island, NY) was added to 500 pmol (total) of siRNAs or mimicRNAs diluted in 2.5 mL of Opti-MEM I Reduced Serum Media (Gibco). To form complexes, Lipofectamine RNAiMAX/RNA solution was incubated for 20 minutes at room temperature then spread in a 100 mm plate. 2.0×10^6 HFFs were added dropwise to the 100 mm plate containing complexes and incubated for 48 h. siRNAs and mimicRNAs were preMir Negative Control #1 (Ambion, Grand Island, NY), Silencer Select Negative Control #1 (Ambion), Silencer Select Validated siRNA Exportin-1 (ID s14937, Ambion), Silencer Select Validated Exportin-1 (ID118206, Ambion), Silencer Select Predesigned siRNA TGS1 (ID s41313, Ambion), siGenome siRNA Exportin-5 (Thermo Scientific, Dharmacon), Silencer Select Validated Exportin-5 (IDs33191, Ambion), Silencer Select Validated RNASEN (Drosha long, IDs26491, Ambion), PIMT (TGS1) siRNA (280 pmol, sc-45875, Santa Cruz), and hsa-miR-34a (ID PM11030, Ambion). For cellular fractionation to determine localization of pri-miRNA in the cytoplasm; 280 pmol of a pool of 3 target-specific 19-25 nt siRNAs against Drosha (sc-44080, Santa Cruz) or siRNA Control was utilized as detailed above to knockdown Drosha in HFF cells.

J. Transient transfection of plasmids

87 μ L of Lipofectamine 2000 (Invitrogen) was diluted in 1.2 mL of Opti-MEM I Reduced Serum Media (Gibco) and incubated at room temperature for 5 minutes. Separately, 11 μ g of plasmid was diluted in 1.2 mL of Opti-MEM I Reduced Serum Media. The Lipofectamine solution and the DNA/OptiMEM solution were mixed and incubated for 20 minutes at room temperature. The Lipofectamine/DNA complexes were added drop-wise to HFFs plated the previous day at a density of 1.5×10^6 cells/100mm plate in antibiotic-free DMEM-10. Antibiotic-free DMEM 10 medium was changed after 4-6 h, and HFFs were incubated overnight at 37°C.

K. Retroviral infection of plasmids

293T cells were transiently transfected with the retroviral plasmid pBABE-puro-Flag-XPO1 or pBABE-puro-Flag-vector and retroviral packaging plasmids (pCL-Eco and VSV-G) using calcium phosphate transfection. Retrovirus was collected after 48 hrs, filtered and supplemented with polybrene (4 μ g/mL) to increase infection efficiency. HFF cells were exposed to retrovirus for 8 hrs and allowed to recover for 48 hrs before use.

L. Luciferase assays

Luciferase plasmids pmirGlo dual-luciferase expression vector (Promega) and derivatives containing the wild-type or mutant Exportin-5 3'UTR (three point mutations in the binding sequence of the miRNA seed region) were transiently transfected into HFFs by using Lipofectamine 2000 as described above. After incubation for 24h, transfected cells were trypsinized and plated at a density of 5.0×10^4 cells/well in 6-well plates in antibiotic-free, serum-free DMEM medium supplemented with MEM non-essential amino acids (for quiescent cells) and 1.0×10^5 cells/well in antibiotic-free, DMEM supplemented with 10% FBS and MEM non-essential amino acids (for proliferating cells) and incubated for 48 h. Luciferase activity was analyzed by using the Dual-Luciferase Reporter Assay System (Promega) per manufacturer's protocol. In short, cells were lysed with Passive Lysis Buffer (250 μ L for quiescent cells and 500 μ L for proliferating cells) by rapid shaking for 30 minutes at room temperature and aliquoted (20 μ L for proliferating cells, 80 μ L for quiescent cells) into 96 well Costar white plate (Corning Incorporated, Corning, NY). Luminescence was measured by using a Modulus II Microplate Luminometer with dual injector system (Turner BioSystems, Sunnyvale, CA). All transfections and luciferase assays were done in triplicate. For microRNA and anti-miR transfections, we transiently transfected 7×10^4 cells/well in a 24 well plate with 17 pmol pmir Negative Control #1 (Ambion) or hsa-miR-34a (PM11030, Ambion) diluted in 87 μ L OptiMEM I Reduced Serum Medium mixed with Lipofectamine RNAiMAX (Life Technologies) as described above. Transfected cells were incubated overnight and then transiently transfected with pmirGlo dual-luciferase expression vector, Exportin-5 3'UTR WT, and Exportin-5 3'UTR MT (0.5 μ g/well) using Lipofectamine 2000 (1 μ L/well) as described above. Cells were then incubated 48 h in antibiotic-free DMEM supplemented with 10% FBS and non-essential amino acids and luciferase assays were performed. Firefly luciferase activity was measured, normalized to Renilla luciferase activity and converted to Relative luciferase units (RLUs). Data average ratios were obtained in each experiment from three replicates; the graph shows the average of two individual experiments with a standard deviation shown as error bars. Two-tailed t-test results are indicated by * for $P < 0.05$ and ** for $P < 0.01$.

M. Cellular Fractionation

Proliferating and quiescent HFF cells were pelleted and immediately processed for cell fractionation to extract either protein or RNA (protocol adapted from Wang, Y. et. al. 2006) (63). Pelleted cells were resuspended in RSB buffer (10 mM Tris, pH 7.5; 0.3 mM sodium chloride; 0.1 mM magnesium chloride), incubated on ice for 5 min, and centrifuged at 1500 rpm for 4 min to re-pellet. Swollen pellets were resuspended in a volume of RSBG40 buffer (10 mM Tris, pH 7.5; 0.3 mM sodium chloride; 0.1 mM magnesium chloride; 10% glycerol; 0.5% NP-40; 100 U/mL RNasin Plus RNase inhibitor (Promega, Madison, WI) and 0.5 mM DTT) at least 4 times the volume of the pellet. Cells were disrupted using an

autoclaved Dounce homogenizer; homogenate was centrifuged at 4000 rpm for 4 min to pellet the nuclear fraction. The cytoplasmic fraction (supernatant) was collected. The pelleted nuclear fraction was resuspended in RSB-G40; 3.3% of sodium deoxycholate and 6.6% of Tween 20 was added to 1/10th of the volume of RSB-G40. Nuclei were re-pelleted by centrifugation at 7000 rpm for three min after five min incubation on ice. The supernatant was collected and added to the cytoplasmic fraction, which was then subjected to Western blot protein analysis, or RNA was extracted using Trizol Reagent (Life Technologies) per manufacturer's instructions and treated with Turbo DNase (Ambion) for 20 minutes at 37°C. RNA concentrations were obtained using a Nanodrop 2000 Spectrophotometer (Thermo Scientific, Wilmington, DE). The nuclear pellet was washed with RSB-G40 and resuspended in NP-40 lysis buffer for protein analysis by Western blot or Trizol for RNA extraction and RT-PCR analysis.

N. Immunoprecipitation of Exportin-1

Immunoprecipitation protocol was adapted from Niranjankumari S *et al.* 2002; Vasudevan S and Steitz JA. 2007; and Vasudevan S *et al.* 2007 (38, 64, 65). Quiescence was induced in HFF cells expressing pBABE-puro-Flag-XPO1 or pBABE-puro-Flag-vector pFlag by serum-starvation (48 hrs). Quiescent HFFs were crosslinked with 1% formaldehyde in PBS for 10 min at RT on a shaker. Glycine was added to a final concentration of 0.125M to quench the reaction for 5 min at RT, shaking occasionally. HFF cells were washed with PBS, scraped and pelleted by centrifugation. Cell pellets lysed with RIPA buffer (50mM tris, pH 7.5, 1% NP-40, 0.5% sodium deoxycholate, 0.05% SDS, 1mM EDTA, 150mM NaCl), and sonicated. Pre-washed anti-Flag M2 Magnetic Bead (20 µL, Sigma-Aldrich) were incubated with 275-300µg of lysates, RNasin Plus RNase Inhibitor (Promega), and 100 µg/mL tRNA (Sigma-Aldrich) while rotating for approximately 4 hrs at 4° C. Beads were washed 6 times with RIPA buffer. RNA/beads were un-crosslinked; briefly, beads were resuspended in 6M urea, 50mM Tris, pH 6.8, and 3mM DTT and incubated for 45 min at 60° C. Proteinase K buffer (50mM Tris-HCL, pH 8.0, 10mM CaCl₂) was added to dilute urea and 75 µg/mL Proteinase K (Fisher Scientific) was added to digest protein. Samples were incubated for 1 hr at 45° C, pelleted and supernatant was collected. RNA was extracted with TRIzol Reagent (Invitrogen) per manufacturer's instructions; GlycoBlue Coprecipitant (Life Technologies) was added.

XIV. Supplemental References

1. **Huber J, Cronshagen U, Kadokura M, Marshallsay C, Wada T, Sekine M, Luhrmann R.** 1998. Snurportin1, an m3G-cap-specific nuclear import receptor with a novel domain structure. *Embo j* **17**:4114-4126.
2. **Pardee AB.** 1974. A restriction point for control of normal animal cell proliferation. *Proc Natl Acad Sci U S A* **71**:1286-1290.

3. **O'Farrell PH.** 2011. Quiescence: early evolutionary origins and universality do not imply uniformity. *Philos Trans R Soc Lond B Biol Sci* **366**:3498-3507.
4. **Coller HA.** 2011. Cell biology. The essence of quiescence. *Science* **334**:1074-1075.
5. **Valcourt JR, Lemons JM, Haley EM, Kojima M, Demuren OO, Coller HA.** 2012. Staying alive: metabolic adaptations to quiescence. *Cell Cycle* **11**:1680-1696.
6. **Coller HA, Sang L, Roberts JM.** 2006. A new description of cellular quiescence. *PLoS Biol* **4**:e83.
7. **Cheung TH, Rando TA.** 2013. Molecular regulation of stem cell quiescence. *Nat Rev Mol Cell Biol* **14**:329-340.
8. **Bohnsack MT, Czaplinski K, Gorlich D.** 2004. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *Rna* **10**:185-191.
9. **Krol J, Loedige I, Filipowicz W.** 2010. The widespread regulation of microRNA biogenesis, function and decay. *Nat Rev Genet* **11**:597-610.
10. **James MA, Lee JH, Klingelhutz AJ.** 2006. HPV16-E6 associated hTERT promoter acetylation is E6AP dependent, increased in later passage cells and enhanced by loss of p300. *Int J Cancer* **119**:1878-1885.
11. **Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, Kim VN.** 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**:415-419.
12. **Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R.** 2004. The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**:235-240.
13. **Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ.** 2004. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**:231-235.
14. **Yi R, Qin Y, Macara IG, Cullen BR.** 2003. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes Dev* **17**:3011-3016.
15. **Hutvagner G, McLachlan J, Pasquinelli AE, Balint E, Tuschl T, Zamore PD.** 2001. A cellular function for the RNA-interference enzyme Dicer in the maturation of the let-7 small temporal RNA. *Science* **293**:834-838.
16. **Grishok A, Pasquinelli AE, Conte D, Li N, Parrish S, Ha I, Baillie DL, Fire A, Ruvkun G, Mello CC.** 2001. Genes and mechanisms related to RNA interference regulate expression of the small temporal RNAs that control *C. elegans* developmental timing. *Cell* **106**:23-34.
17. **Knight SW, Bass BL.** 2001. A role for the RNase III enzyme DCR-1 in RNA interference and germ line development in *Caenorhabditis elegans*. *Science* **293**:2269-2271.
18. **Han J, Lee Y, Yeom KH, Kim YK, Jin H, Kim VN.** 2004. The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev* **18**:3016-3027.
19. **Lian SL, Li S, Abadal GX, Pauley BA, Fritzler MJ, Chan EK.** 2009. The C-terminal half of human Ago2 binds to multiple GW-rich regions of GW182 and requires GW182 to mediate silencing. *RNA* **15**:804-813.
20. **Hammond SM, Boettcher S, Caudy AA, Kobayashi R, Hannon GJ.** 2001. Argonaute2, a link between genetic and biochemical analyses of RNAi. *Science* **293**:1146-1150.
21. **Mourelatos Z, Dostie J, Paushkin S, Sharma A, Charroux B, Abel L, Rappsilber J, Mann M, Dreyfuss G.** 2002. miRNPs: a novel class of ribonucleoproteins containing numerous microRNAs. *Genes Dev* **16**:720-728.
22. **Ruby JG, Jan CH, Bartel DP.** 2007. Intronic microRNA precursors that bypass Drosha processing. *Nature* **448**:83-86.
23. **Cazalla D, Xie M, Steitz JA.** 2011. A primate herpesvirus uses the integrator complex to generate viral microRNAs. *Mol Cell* **43**:982-992.
24. **Havens MA, Reich AA, Duelli DM, Hastings ML.** 2012. Biogenesis of mammalian microRNAs by a non-canonical processing pathway. *Nucleic Acids Res* **40**:4626-4640.
25. **Cifuentes D, Xue H, Taylor DW, Patnode H, Mishima Y, Cheloufi S, Ma E, Mane S, Hannon GJ, Lawson ND, Wolfe SA, Giraldez AJ.** 2010. A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* **328**:1694-1698.

26. **Cheloufi S, Dos Santos CO, Chong MM, Hannon GJ.** 2010. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* **465**:584-589.
27. **Shapiro JS, Langlois RA, Pham AM, Tenoever BR.** 2012. Evidence for a cytoplasmic microprocessor of pri-miRNAs. *Rna* **18**:1338-1346.
28. **Xie M, Li M, Vilborg A, Lee N, Shu MD, Yartseva V, Sestan N, Steitz JA.** 2013. Mammalian 5'-capped microRNA precursors that generate a single microRNA. *Cell* **155**:1568-1580.
29. **Hutten S, Kehlenbach RH.** 2007. CRM1-mediated nuclear export: to the pore and beyond. *Trends Cell Biol* **17**:193-201.
30. **Ohno M, Segref A, Bachi A, Wilm M, Mattaj JW.** 2000. PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation. *Cell* **101**:187-198.
31. **Palacios I, Hetzer M, Adam SA, Mattaj JW.** 1997. Nuclear import of U snRNPs requires importin beta. *EMBO J* **16**:6783-6792.
32. **Kiss T.** 2004. Biogenesis of small nuclear RNPs. *J Cell Sci* **117**:5949-5951.
33. **Boulon S, Verheggen C, Jady BE, Girard C, Pescia C, Paul C, Ospina JK, Kiss T, Matera AG, Bordonne R, Bertrand E.** 2004. PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. *Mol Cell* **16**:777-787.
34. **Yedavalli VS, Jeang KT.** 2010. Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs. *Proc Natl Acad Sci U S A* **107**:14787-14792.
35. **Medina R, Zaidi SK, Liu CG, Stein JL, van Wijnen AJ, Croce CM, Stein GS.** 2008. MicroRNAs 221 and 222 bypass quiescence and compromise cell survival. *Cancer Res* **68**:2773-2780.
36. **Suh EJ, Remillard MY, Legesse-Miller A, Johnson EL, Lemons JM, Chapman TR, Forman JJ, Kojima M, Silberman ES, Collier HA.** 2012. A microRNA network regulates proliferative timing and extracellular matrix synthesis during cellular quiescence in fibroblasts. *Genome Biol* **13**:R121.
37. **Martinez I, Cazalla D, Almstead LL, Steitz JA, DiMaio D.** 2011. miR-29 and miR-30 regulate B-Myb expression during cellular senescence. *Proc Natl Acad Sci U S A* **108**:522-527.
38. **Vasudevan S, Tong Y, Steitz JA.** 2007. Switching from repression to activation: microRNAs can up-regulate translation. *Science* **318**:1931-1934.
39. **Iwasaki YW, Kiga K, Kayo H, Fukuda-Yuzawa Y, Weise J, Inada T, Tomita M, Ishihama Y, Fukao T.** 2013. Global microRNA elevation by inducible Exportin 5 regulates cell cycle entry. *RNA* **19**:490-497.
40. **Link S, Grund SE, Diederichs S.** 2016. Alternative splicing affects the subcellular localization of Drosha. *Nucleic Acids Res* **44**:5330-5343.
41. **Rissland OS, Hong SJ, Bartel DP.** 2011. MicroRNA destabilization enables dynamic regulation of the miR-16 family in response to cell-cycle changes. *Mol Cell* **43**:993-1004.
42. **Bussing I, Yang JS, Lai EC, Grosshans H.** 2010. The nuclear export receptor XPO-1 supports primary miRNA processing in *C. elegans* and *Drosophila*. *EMBO J* **29**:1830-1839.
43. **Castanotto D, Lingeman R, Riggs AD, Rossi JJ.** 2009. CRM1 mediates nuclear-cytoplasmic shuttling of mature microRNAs. *Proc Natl Acad Sci U S A* **106**:21655-21659.
44. **Luhrmann R, Appel B, Bringmann P, Rinke J, Reuter R, Rothe S, Bald R.** 1982. Isolation and characterization of rabbit anti-m3 2,2,7G antibodies. *Nucleic Acids Res* **10**:7103-7113.
45. **Girard C, Verheggen C, Neel H, Cammas A, Vagner S, Soret J, Bertrand E, Bordonne R.** 2008. Characterization of a short isoform of human Tgs1 hypermethylase associating with small nucleolar ribonucleoprotein core proteins and produced by limited proteolytic processing. *J Biol Chem* **283**:2060-2069.
46. **Chong MM, Zhang G, Cheloufi S, Neubert TA, Hannon GJ, Littman DR.** 2010. Canonical and alternate functions of the microRNA biogenesis machinery. *Genes Dev* **24**:1951-1960.
47. **Kawai S, Amano A.** 2012. BRCA1 regulates microRNA biogenesis via the DROSHA microprocessor complex. *J Cell Biol* **197**:201-208.

48. **Kim YK, Kim B, Kim VN.** 2016. Re-evaluation of the roles of DROSHA, Export in 5, and DICER in microRNA biogenesis. *Proc Natl Acad Sci U S A* **113**:E1881-1889.
49. **Ha M, Kim VN.** 2014. Regulation of microRNA biogenesis. *Nat Rev Mol Cell Biol* **15**:509-524.
50. **Legesse-Miller A, Raitman I, Haley EM, Liao A, Sun LL, Wang DJ, Krishnan N, Lemons JM, Suh EJ, Johnson EL, Lund BA, Collier HA.** 2012. Quiescent fibroblasts are protected from proteasome inhibition-mediated toxicity. *Mol Biol Cell* **23**:3566-3581.
51. **Gibbins D, Mostowy S, Jay F, Schwab Y, Cossart P, Voinnet O.** 2012. Selective autophagy degrades DICER and AGO2 and regulates miRNA activity. *Nat Cell Biol* **14**:1314-1321.
52. **Bitto A, Lerner C, Torres C, Roell M, Malaguti M, Perez V, Lorenzini A, Hrelia S, Ikeno Y, Matzko ME, McCarter R, Sell C.** 2010. Long-term IGF-I exposure decreases autophagy and cell viability. *PLoS One* **5**:e12592.
53. **Verheggen C, Bertrand E.** 2012. CRM1 plays a nuclear role in transporting snoRNPs to nucleoli in higher eukaryotes. *Nucleus* **3**:132-137.
54. **Karamysheva Z, Diaz-Martinez LA, Warrington R, Yu H.** 2015. Graded requirement for the spliceosome in cell cycle progression. *Cell Cycle* **14**:1873-1883.
55. **Shapiro JS, Schmid S, Aguado LC, Sabin LR, Yasunaga A, Shim JV, Sachs D, Cherry S, tenOever BR.** 2014. Drosha as an interferon-independent antiviral factor. *Proc Natl Acad Sci U S A* **111**:7108-7113.
56. **He L, He X, Lim LP, de Stanchina E, Xuan Z, Liang Y, Xue W, Zender L, Magnus J, Ridzon D, Jackson AL, Linsley PS, Chen C, Lowe SW, Cleary MA, Hannon GJ.** 2007. A microRNA component of the p53 tumour suppressor network. *Nature* **447**:1130-1134.
57. **Chen F, Hu SJ.** 2012. Effect of microRNA-34a in cell cycle, differentiation, and apoptosis: a review. *J Biochem Mol Toxicol* **26**:79-86.
58. **Bader AG.** 2012. miR-34 - a microRNA replacement therapy is headed to the clinic. *Front Genet* **3**:120.
59. **Li P, Liu Y, Yi B, Wang G, You X, Zhao X, Summer R, Qin Y, Sun J.** 2013. MicroRNA-638 is highly expressed in human vascular smooth muscle cells and inhibits PDGF-BB-induced cell proliferation and migration through targeting orphan nuclear receptor NOR1. *Cardiovasc Res* **99**:185-193.
60. **Jia L, Wu J, Zhang L, Chen J, Zhong D, Xu S, Xie C, Cai J.** 2013. Restoration of miR-1228* expression suppresses epithelial-mesenchymal transition in gastric cancer. *PLoS One* **8**:e58637.
61. **Lund E, Guttinger S, Calado A, Dahlberg JE, Kutay U.** 2004. Nuclear export of microRNA precursors. *Science* **303**:95-98.
62. **Jia Y, Viswakarma N, Crawford SE, Sarkar J, Sambasiva Rao M, Karpus WJ, Kanwar YS, Zhu YJ, Reddy JK.** 2012. Early embryonic lethality of mice with disrupted transcription cofactor PIMT/NCOA6IP/Tgs1 gene. *Mech Dev* **129**:193-207.
63. **Wang Y, Zhu W, Levy DE.** 2006. Nuclear and cytoplasmic mRNA quantification by SYBR green based real-time RT-PCR. *Methods* **39**:356-362.
64. **Vasudevan S, Steitz JA.** 2007. AU-rich-element-mediated upregulation of translation by FXR1 and Argonaute 2. *Cell* **128**:1105-1118.
65. **Niranjanakumari S, Lasda E, Brazas R, Garcia-Blanco MA.** 2002. Reversible cross-linking combined with immunoprecipitation to study RNA-protein interactions in vivo. *Methods* **26**:182-190.

Chapter 4

Overall Discussion and Future Directions

Traditionally, biomedical research was conducted based on the central dogma of molecular biology, which is that DNA contains the genomic information, RNA is the "messenger" of this information, and proteins are the final product. In 2007, the Encyclopedia of DNA Elements (ENCODE) showed that more than 80% of the human genome is transcribed, but less than 3% of these molecules become proteins(1). These numerous newly discovered transcribed RNAs are classified as non-coding RNAs (ncRNAs) and have now evoked researchers to identify novel ncRNA regulators of cellular pathways besides the 30,000 protein-coding genes encoded by the human genome. A single ncRNA can regulate many different genes involved in many different cellular pathways, so it is critical to further study the role(s) of these individual ncRNAs in more detail.

NcRNA expressions are altered by a variety of cellular processes that are involved with the progression to cancer, such as those occurring after infection with oncogenic viruses (e.g. human papillomavirus)(2, 3) or specifically involved with cell cycle regulation, such as cellular growth arrest (e.g. quiescence) (Chapter 3)(4). Our group's overall goal is to continue to identify early molecular changes in ncRNAs that occur in response to cellular stresses and contribute to carcinogenesis. In the context of cancer, traditionally, ncRNA expression changes are identified by comparing established tumors to normal tissues. One of the specific goals of our group is to continue to elucidate *early* molecular events that cause initial dysregulation but are also critical enough to be *maintained* throughout carcinogenesis.

Oncogenic viruses, such as human papillomavirus (HPV), are useful models to study early events in cancer because they are single exogenous agents involved in the initial steps of carcinogenesis. HPV infection can regulate some unique cellular pathways, but also some analogous pathways other carcinogenic factors utilize. Thus, identifying early molecular events that are critical for progression to carcinogenesis in HPV-associated cancers could potentially be applicable to other cancers and would add to the marketability of these involved genes as therapeutic targets or diagnostic markers.

Additionally, it is necessary to continue to study HPV, as it continues to be linked to development of 4.5% of total cancer incidences worldwide. Certain HPV-associated cancers, such as the oropharyngeal

subset of head and neck squamous cell carcinoma (HNSCC), are on the rise even in industrialized countries(5). In principle, HPV-associated cancer should be preventable as there are prophylactic vaccinations available to prevent infection with strains commonly associated with oncogenesis. However, more than two-thirds of the cervical cancer incidences occur in less-developed countries(6), which are limited by financial and logistical issues for access to the vaccines(7). Worldwide, as of October 2014, only a small minority (7.5%) of 10 to 20-year-old females are estimated to have been administered at least one shot of an HPV vaccine(8), and this statistic does not guarantee that all of these patients completed the full administration of shots required for full efficacy. Additionally, since the HPV vaccinations are prophylactic, they do not protect those who have already been exposed to the virus. According to the CDC, "HPV is so common that nearly all men and women get exposed to the virus at some point in their lives" (<https://www.cdc.gov/std/hpv/stats.htm>). Thus, these many individuals who already have been exposed to oncogenic strains of the virus risk persistent infection contributing to carcinogenesis, and new cancer therapies need to be developed.

HPV regulates a variety of downstream genes to contribute to carcinogenesis(9, 10). Our group recently revealed novel coding genes (including LAT1 and NAP1L3) that are differentially expressed in HPV-positive compared to HPV-negative cervical cancer lines and normal cervical keratinocytes. Future studies will be conducted to elucidate downstream effects of their up-regulation as well as potential use as biomarkers associated with the development of cervical cancer. Different types of ncRNAs [microRNAs (miRNAs) and long non-coding RNAs (lncRNAs)] also exhibit global differential expression when comparing cervical cancer patients (expressing the full HPV genome) to normal matched cervical tissues (2, 3).

Notably, lncRNAs are a more recent field of interest and many have not been fully characterized. There is not a clear, set standard collection of lncRNAs that all research groups analyze; certain lncRNAs might not be analyzed in every dataset so it can be difficult to compare between different groups' studies. In the initial stages of the project presented in Chapter 2, MD Anderson's The Atlas of ncRNA in Cancer (TANRIC) database (http://ibl.mdanderson.org/tanric/_design/basic/index.html) was the most user-friendly, open access resource for extracting information from large high-throughput RNA sequencing (RNA-seq) datasets showing lncRNAs expression changes involved in cancer. More recently, additional resources are available such as The Cancer LncRNome Atlas (TCLA) (<http://tcla.fcgportal.org/>) database hosted by University of Pennsylvania.

Two specific oncogenic components of the HPV genome, E6 and E7, affect two main tumor suppressors p53(11) and Rb(12), respectively, however they have been found to regulate a plethora of genes involved in cancer progression(10, 13). Since the HPV oncogenes regulate pathways commonly altered in cancer (even in HPV-negative cancers), looking at downstream pathways regulated by the HPV oncogenes can also be applicable to other types of cancers. In addition to coding gene regulation specifically by HPV oncogenes, certain studies have shown ncRNA regulation. Martinez et al. showed high-risk HPV-16 E6 regulation of a miRNA (miR-218) that regulates a coding gene (LAMB3) to potentially increase viral infection of surrounding tissue and contribute to carcinogenesis (14). Additionally, a couple previous studies showed that the HPV-16 E6/E7 oncogenes regulate the expression of lncRNA MALAT1(15) and CCEPR(16) to contribute to cervical cancer. Our group recently showed that HPV E6 expression alone causes differential expression alterations in additional host lncRNAs (Chapter 2, Figure 1A and Supplementary Table S1). Future studies elucidating the mechanism(s) of these HPV E6-regulated lncRNAs could contribute significantly to the cancer field.

Another interesting future route to study HPV E6 regulation of lncRNAs is to evaluate if any lncRNAs are directly bound to the HPV-16 E6 oncoprotein. A variety of proteins are shown to directly interact with HPV-16 E6(17), but enhancement of knowledge on ncRNA E6 binding partners could reveal alternative downstream regulators of carcinogenesis. Since HPV-16 E6 is localized to the nucleus(18), it is predicted the bound ncRNAs will be nuclear. The primary goal of this direction would be to identify host lncRNAs that are bound to E6, however, with advanced high-throughput RNA sequencing technologies it would be possible to identify novel host and viral ncRNAs bound to HPV. The Martinez lab has unpublished data showing HPV produces viral circular RNAs that alter different hallmarks of cancer (proliferation and apoptosis) (data not shown). Our lab has not yet elucidated if this is directly or indirectly by HPV-16 E6 or the presence of the full genome. Preliminary experiments were conducted to attempt to transfect a Flag-HPV-16 E6 construct into cervical cells and immunoprecipitate out by an antibody against Flag. There are well-known technical difficulties in the HPV field when trying to detect HPV-16 E6 because good quality antibodies are lacking and it has such a low molecular weight (~21 kDa) that is difficult by Western Blot analysis.

Our lab is particularly interested in researching the host lncRNA FAM83H-AS1, which our group identified to be HPV-16 E6 regulated in cervical cells (Chapter 2, Figure 3A-C, Supplementary Figure S4A,B), and its importance in cancer initiation and progression. Most lncRNAs are tissue specific(19). FAM83H-AS1 is classified as an "onco-lncRNA", which describes the few lncRNAs that have been determined to be altered in many different cancer tissue lineages unrelated to HPV infection. This is an

important feature because if a lncRNA is affected in multiple types of cancers, it suggests its universal importance in essential cellular pathways involved in carcinogenesis. The study of these onco-lncRNAs could increase the ability to develop potential diagnostic, prognostic, or therapeutic applications for several types of cancers(20). FAM83H-AS1 expression was shown to be up-regulated in the following cancers and high expression of FAM83H-AS1 in each of these cancers correlates with poor overall survival: breast(20, 21), colorectal(20, 22, 23), lung(20, 24), pancreatic(25), brain(26), and cervical (Chapter 2, Figure 6) (Figure 1). Therefore, further elucidating the FAM83H-AS1 in HPV-related cancers could be beneficial to a broad range of cancer patients.

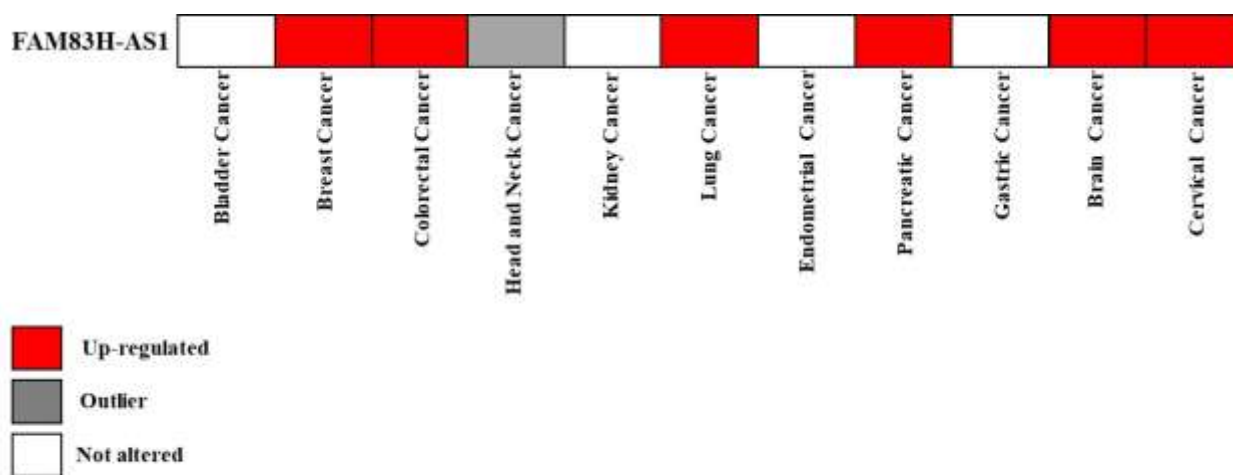


Figure 1: Heat map of FAM83H-AS1 expression alterations in multiple cancer types.
Adapted from: Cabanski et al., *RNA Biology*, 2015.

Our group showed for the first time that FAM83H-AS1 expression is altered at an early stage of HPV infection in cervical cells (Chapter 2, Figure 2A), and this alteration is maintained throughout later stages of cervical carcinogenesis (Chapter 2, Figure 2B and Figure 6A,B). These results could be further strengthened in the future with access to larger sample sizes of normal, pre-malignant (cervical intraepithelial neoplasia), and cancerous human patient tissues to analyze FAM83H-AS1 expression differences. These data would contribute to the conclusion that initial up-regulation of FAM83H-AS1 is an early event similar to HPV infection itself. This increases its applicability for utilization as a biomarker to predict patients with high FAM83H-AS1 expression changes early on with HPV infection who will likely progress to cervical carcinogenesis.

Our group analyzed FAM83H-AS1 expression after persistent HPV infection by extracting data from TANRIC to show high expression of FAM83H-AS1 in cervical squamous cell carcinoma and endocervical adenocarcinoma (CESC) patients and that high expression of FAM83H-AS1 correlated with poor overall survival (Chapter 2, Figure 6B,C). HPV is detected in almost all cervical cancers. HPV-16 is

the most prevalent strain (50%), HPV-18 the second most common (15%), and there are many other strains (-31, -33, -35, -39, -45, -51, -52, -56, -58, -59, and 68) classified as carcinogenic(27). Dividing TANRIC CESC patients into subgroups by strain could increase the impact if it is found that FAM83H-AS1 expression is altered in more than just HPV-16 positive samples. The original article that classified FAM83H-AS1 as an onco-lncRNA showed that, in their analysis, its expression in head and neck cancer is an outlier, where "outlier" is defined as a subset that had an alteration in expression(20). It is speculated that if these head and neck samples were divided into HPV-positive and HPV-negative subsets that there could have been a correlation determined. Further data mining from TANRIC of other cancer types, and segregation of samples based on HPV status, could be utilized to enhance the idea that FAM83H-AS1 is up-regulated in an HPV-dependent manner in certain cancer types.

HNSCC is a significant cancer type to study further as recently it was identified that the incidence of HPV-positive oropharyngeal cancer has surpassed cervical cancer in the United States(5). Our group showed that FAM83H-AS1 expression was also shown to be higher in HPV-positive versus HPV-negative HNSCC cell lines (Chapter 2, Figure 2C). Future access to human HPV-positive oropharyngeal HNSCC tissue samples as well as functional assays after FAM83H-AS1 knockdown in HNSCC cell lines would be beneficial to strengthen the HPV regulation of FAM83H-AS1 in HNSCC samples.

Overall, since individual knockdown of FAM83H-AS1 showed positive responses desired for cervical cancer cells (decreased proliferation and migration, and increased apoptosis) (Chapter 2, Figure 5 and Supplementary Figure S7), FAM83H-AS1 expression could potentially be targeted in either early stages of HPV infection (when the increase in expression is detected such that it could potentially prevent progression to carcinogenesis) or in later stages of cervical cancer.

Genes can be alternatively spliced under cellular stresses, which can contribute to carcinogenesis, and specifically HPV gene expression is controlled by alternative splicing to regulate its viral life cycle(28). Our group has obtained unpublished data showing alternative splicing events upon HPV infection can produce a viral circular RNA that regulates proliferation in cervical cancer cells (data not shown). Our lab plans to continue to elucidate alternatively spliced host and virus-produced ncRNAs critical for the initiation and progression to cancer.

Certain lncRNAs are predicted to have multiple isoforms according to online databases such as LNCipedia (lncipedia.org). We plan to fully sequence the specific functional isoform(s) of FAM83H-AS1 that is important in cervical cancer. Northern Blot probes, developed by *in vitro* transcription, will be used

to detect the sizes of the FAM83H-AS1 isoforms. Then, TANRIC's RNA sequencing data of FAM83H-AS1 could be extracted and bioinformatics analysis conducted to determine the frequency of reads at certain regions of RNA transcript. The sequences of these regions could be extracted from UCSC Genome Browser and used to design sets of primers to confirm (via RT-PCR) the presence of these certain RNA segments and map the entire functional RNA sequence. Once the functional FAM83H-AS1 isoform(s) is characterized in cervical cells, the presence and functionality of this specific isoform(s) in other cancer lineages could be determined to increase the universal potential of this lncRNA as a therapeutic target and/or diagnostic marker. Sequencing the functional isoform of FAM83H-AS1 would enable multiple future studies described in the paragraphs below.

Elucidation of upstream regulators of FAM83H-AS1 could determine ways to control and/or predict the lncRNA's expression changes. Our group showed that FAM83H-AS1 expression is regulated by HPV-16 E6 in cervical cells (Chapter 2, Figure 3A,B,C, Supplementary Figure S4A,B). This regulation of FAM83H-AS1 by HPV could be characterized further. Reporter assays, using a luciferase reporter vector containing the promoter region of FAM83H-AS1, could be transfected into HPV-positive and HPV-negative cells with or without siRNAs against HPV oncogenes E6 and/or E7. It is expected that there will be changes in luciferase activity in HPV-positive cell lines after the repression of the HPV oncogenes, suggesting a regulation of FAM83H-AS1 at the transcriptional level. If the hypothesis is correct, different luciferase constructs with specific mutations in the FAM83H-AS1 promoter could be used to identify the regulatory region in the genomic sequence.

HPV-16 E6 contributes to the degradation of the tumor suppressor p53, which alters the expression of many p53-regulated genes and affects critical cellular pathways (e.g. proliferation, DNA repair, and apoptosis)(9, 29) to contribute to carcinogenesis. Interestingly, FAM83H-AS1 was determined to be regulated in a p53-independent manner in cervical keratinocytes (Chapter 2, Figure 3D, Supplementary Figure S4C). Since HPV-16 E6 interacts directly with p300(17, 30, 31), and that there are three p300 binding sites in the predicted promoter region of FAM83H-AS1 (Chapter 2, Figure 3E) as determined by using the transcription factor binding site prediction software ConTra v3(32) we analyzed FAM83H-AS1 regulation by p300. The transcription activator p300 was identified as a regulator of FAM83H-AS1 expression in cervical keratinocytes (Chapter 2, Figure 3F, Supplementary Figure S4D). To further strengthen this, an inducible p300 knockdown construct could be stably transduced into cervical cancer cells and FAM83H-AS1 expression could be monitored by a dose-dependent progressive knockdown of p300. Reciprocally, p300 could be added to cervical keratinocytes (likely exhibiting low baseline expression of p300) and FAM83H-AS1 expression could be monitored; it would be expected that

FAM83H-AS1 expression would increase with p300 expression. A reporter assay, using a luciferase reporter vector containing the promoter-region of FAM83H-AS1, could be used to determine direct binding between p300 and FAM83H-AS1. The expression of FAM83H-AS1 could be monitored between a luciferase reporter vector containing the wild-type promoter as well as different constructs containing mutations in each of the individual p300 binding sites. Additionally, to confirm direct binding, chromatin immunoprecipitation (ChIP) with a p300 antibody could be used to determine direct binding of p300 to the FAM83H-AS1 promoter region of the genome.

Since other groups previously showed FAM83H-AS1 expression is altered in cancers not regulated by HPV expression(20-26), HPV is not the only potential driver of increased FAM83H-AS1 expression. It is speculated that HPV-16 E6 could be regulating a downstream factor, and this downstream factor is more directly responsible for regulating FAM83H-AS1 expression. A group previously analyzed specific proteins that are bound to the E6 oncoproteins expressed by different high-risk strains of HPV. Interestingly, of the high-risk strains that were analyzed, the only high-risk strain E6 oncoprotein that bound transcription activator p300 was HPV-16(17) (Figure 2). Our group performed preliminary analysis of FAM83H-AS1 expression in an HPV-18 positive cervical carcinoma cell line (HeLa) and found that FAM83H-AS1 is

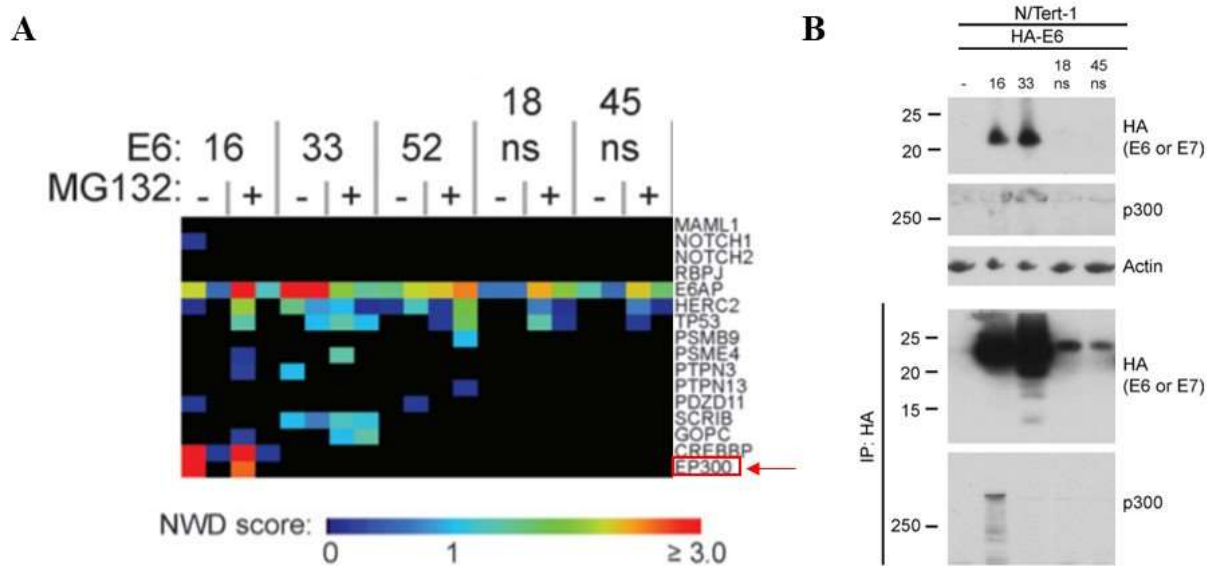


Figure 2: HPV-16 was the only high-risk strain analyzed that demonstrated E6 oncoprotein binding to p300. **A)** Heat map of interactions observed after immunoprecipitation of E6 oncoprotein and mass spectroscopy analysis of bound proteins. **B)** Western blot analysis to confirm that p300 immunoprecipitated out with HA-tagged E6 proteins. Adapted from: Cabanski et al., RNA Biology, 2015.

expressed at similar levels to normal cervical keratinocytes (Chapter 2, Supplementary Figure S4E). Therefore, our group speculates that HPV-16 E6 regulates FAM83H-AS1 indirectly in cervical cancer, and p300 is directly regulating FAM83H-AS1 expression in HPV-16 positive cervical cancer. Furthermore, other types of cancers where FAM83H-AS1 is found overexpressed also showed overexpression of the transcription factor p300(33-37).

Many antisense lncRNAs can share the same promoter as nearby coding genes and their expressions correlate(38, 39). LncRNA FAM83H-AS1 is located near the coding gene FAM83H in the genome and transcribed antisense to the coding gene. Thus, it was hypothesized that they share the same promoter. FAM83H is up-regulated in many different types of cancer(40), and HPV-16 E6 up-regulates FAM83H as well as FAM83H-AS1 expression (Chapter 2, Figure 3E, Supplementary Figure S5B). Thus, it can be speculated that the lncRNA and coding gene share the same promoter. To confirm this, the sense and antisense sequences could be cloned independently in a luciferase reporter vector and their activity verified. Additionally, knocking down p300 and analyzing for changes in FAM83H expression would provide evidence that FAM83H and FAM83H-AS1 share the same promoter region.

In addition to identifying upstream regulators of FAM83H-AS1, it is critical to identify specific downstream effects of dysregulated FAM83H-AS1 expression. Previous publications showed that FAM83H-AS1 regulates CDKN1A in glioma cells(26), Notch signaling in colorectal cancer(22), and MET/EGFR signaling in lung cancer(24). These publications specifically looked at FAM83H-AS1 downstream regulation in single cancer types, thus it could be determined if FAM83H-AS1 functions similarly in other cancer types including cervical cancers.

In the ncRNA field, it is understood that the stoichiometry between a ncRNA and its targets can predict the chances of interactions and consequential downstream effects. NcRNA expression is usually quantified using techniques such as qRT-PCR analysis that show relative expression. However, it is important to determine absolute expression of RNAs in cells. To do this, *in vitro* transcription followed by qRT-PCR analysis can be used. We believe, however, that FAM83H-AS1 is expressed at a reasonable functional level, as its RPKM value in our RNAseq data is similar to the RPKM values of known functional lncRNAs (Chapter 2, Supplementary Table S2).

Determining the localization of a lncRNA can aid in elucidating the downstream function of the lncRNA. For example, if a lncRNA is nuclear it can be speculated that it regulates at the transcriptional level. Other groups have shown the localization of FAM83H-AS1 in the nucleus in lung and breast cancer

cells(24, 26), and our group recently found FAM83H-AS1 in the nucleus of cervical cancer cells (as seen in Chapter 2 with cellular fractionation in Figure 4A,B). Fully sequencing FAM83H-AS1 would provide information necessary to design RNA fluorescent in situ hybridization (FISH) probes that could be used to detect the localization of lncRNAs and strengthen the finding that FAM83H-AS1 is localized in the nucleus. RNA FISH probes could hybridize with cervical cells as well as cervical tissue microarrays to analyze FAM83H-AS1 expression levels in human cervical samples at early stage infection (cervical intraepithelial neoplasia) through late stage carcinogenesis. Successful design of RNA FISH probes for FAM83H-AS1 would be beneficial to study other cancer types as well, as RNA FISH probes have not been used to detect FAM83H-AS1 in any cellular context to our knowledge.

Many nuclear lncRNAs can function in cis(41). Since FAM83H-AS1 is localized in the nucleus in cervical cancer cells (Chapter 2, Figure 4A,B), it was initially hypothesized that FAM83H-AS1 regulates the nearby coding gene FAM83H in cis. Knockdown of FAM83H-AS1 lncRNA does not appear to regulate FAM83H expression (Chapter 2, Supplementary Figure S5C), thus our group speculates that FAM83H-AS1 does not regulate FAM83H in cis in our system. FAM83H-AS1 is likely functioning in trans in our model, but it does not eliminate the possibility that FAM83H-AS1 could function in cis in other cancer types or cellular stress situations (e.g. serum starvation and hypoxia).

Determining RNA and protein expressions that are affected by FAM83H-AS1 would potentially allow us to further narrow down the lncRNA mechanism of action. Using samples with efficient knockdown of FAM83H-AS1 in cervical cancer cells, it is proposed to conduct RNA high-throughput sequencing and mass spectroscopy to determine specific direct and/or indirect downstream effectors of FAM83H-AS1 that are contributing to the observed carcinogenic effects. A previous study analyzed differential gene expression changes with two different siRNAs against FAM83H-AS1 in a pancreatic cancer cell line, Aspc1(25). The theoretical dataset in cervical cancer could be compared to the pancreatic cancer dataset to identify similar molecular regulation by FAM83H-AS1 between cancer types. To elucidate direct binding partners, FAM83H-AS1 could be pulled out by a biotinylated RNA probe to identify its direct RNA/DNA (high-throughput RNA sequencing) and protein (mass spectroscopy) binding partners. A previous publication showed that in glioma cells, FAM83H-AS1 binds to EZH2(26), thus this would be a binding target we would look for in cervical cancer cells, as EZH2 regulates many genes at the transcriptional level by DNA methylation. Both cytoplasmic(42) and nuclear(43) lncRNAs can bind to miRNAs to act as competing endogenous RNAs. To verify potential binding interactions between FAM83H-AS1 and miRNAs to suggest a post-transcriptional regulation of miRNAs, crosslinking immunoprecipitation (CLIP) assays using Argonaute antibodies (that is the protein of the RNA induced silencing complex that miRNAs

are loaded into) could be conducted. Once the analysis is narrowed down to critical lncRNAs, Target scan could be used to determine potential FAM83H-AS1 miRNA binding sites. To confirm direct miRNA binding to lncRNA, luciferase reporter experiments could be conducted by cloning FAM83H-AS1 regions with potential miRNA binding sites to a luciferase reporter vector. Subsequently, FAM83H-AS1 activity could be measured by using siRNAs against specific miRNAs and then mutating the potential binding sites in these constructs. Other FAM83H-AS1 binding partners are yet to be determined, thus this experiment would provide novel information.

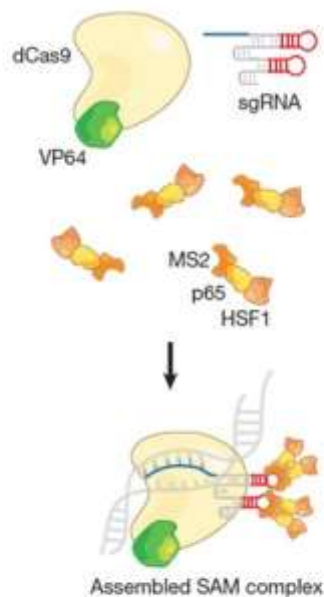


Figure 3: Schematic of the SAM complex. Complex is composed of three components: dCas9-VP64; MS2-p65-HSF1; sgRNA. Adapted from: Konermann et al., *Nature*, 2015.

In addition to specific genes regulated by FAM83H-AS1, we are interested in monitoring functional changes with alterations in FAM83H-AS1 expression. Knockdown of FAM83H-AS1 in cervical cancer cells resulted in decreased proliferation and migration, increased cells in S-phase, and increased apoptosis (Chapter 2, Figure 6 and Supplementary Figure S7). An ideal next step would be to express the FAM83H-AS1 in cells with low baseline FAM83H-AS1 expression and monitor for functional changes. The hypothesis for this future direction is that expression of FAM83H-AS1 would increase proliferation and migration. There was a group

recently that cloned the PCR-amplified complementary DNA (cDNA) sequence that encoded FAM83H-AS1 into an expression vector and transfected this into cells to monitor functional changes(26). However, to ensure expression of the entire functional lncRNA, we propose cloning the full FAM83H-AS1 sequence into a retroviral plasmid (e.g. PLPCX) or use the lentiviral CRISPR-Cas9 mutant technology known as SAM complex (Figure 3). This system requires sequentially expressing the lentiviral plasmids (Addgene) MS2-P65-HSF1_Hygro (Plasmid #1), dCAS9-VP64_Blast (Plasmid #2), and sgRNA(MS2)_Zeo (Plasmid #3)(44). Briefly, Plasmid #1 expresses the MS2 protein that binds to p65 and heat-shock factor 1 (HSF1) proteins that enhance transcription factor signals. Plasmid #2 contains a mutant Cas9 (dCas9) that inhibits its endonuclease activity, and is bound to the transcription factor VP64. Plasmid #3 was cloned such that it would specifically target a predicted FAM83H-AS1 promoter site. Preliminary data was obtained in the lab showing successful stable transfection of Plasmids #1 (MS2 mRNA), #2 (dCas9 mRNA), and #3 (Zeocin

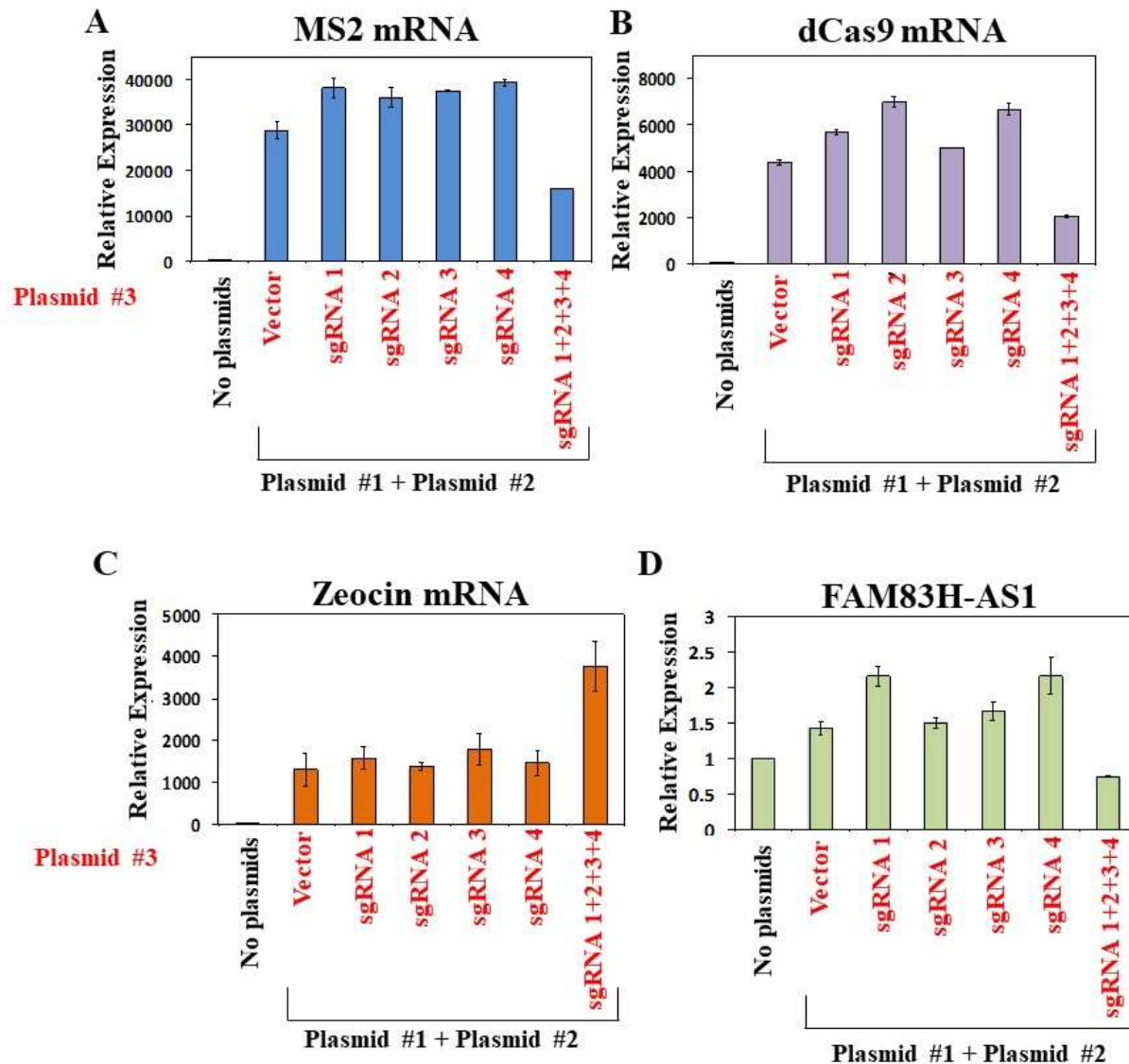


Figure 4: CRISPR Cas9 mutant technology to over-express FAM83H-AS1 in C-33A cell line. Confirmation that A) Plasmids #1, B) Plasmid #2, and C) and Plasmid #3 are stably expressing desired genes. D) Plasmid #3 sgRNA designed to target FAM83H-AS1 promoter did not induce FAM83H-AS1.

mRNA) in an HPV-negative cervical carcinoma cell line (C-33A) and immortalized keratinocytes (HaCat). However, there was not significant up-regulation of FAM83H-AS1 observed (Figure 4). After further analysis, it was determined that the predicted promoter sequence inserted into Plasmid #3 was not the region that we currently believe is the predicted promoter site (region containing p300 binding sites). This would explain the lack of increased FAM83H-AS1 expression observed in Figure 4. A future direction is to re-design sgRNAs to target the more characterized FAM83H-AS1 promoter site and over-express FAM83H-

AS1 in cervical cells. This system would be useful to study the effects of increased FAM83H-AS1 expression in other cancers as well.

In the knockdown experiments described in Chapter 2, FAM83H-AS1 expression was transiently knocked down. It would be required to develop a system to stably knockdown FAM83H-AS1 for future *in vivo* experiments and long-term *in vitro* experiments (e.g. colony formation assays); it is important to note that the transient knockdown of FAM83H-AS1 was maintained for at least 120 hours (Chapter 2, Supplementary Figure S6). In developing a system to stably knockdown FAM83H-AS1, it would also be ideal to develop a stable population from a single cell clone population to ensure that all cells have knockdown of FAM83H-AS1 and there is a lack of non-transfected cell background.

It is clear that a significant portion of the cells had knockdown of FAM83H-AS1, as we saw significant functional changes, however the transfection efficiency was not determined. In the future, a system could be developed to express a fluorescent protein (e.g. GFP) to confirm siRNA successfully transfected into cell nuclei. This would allow for either visualization under fluorescent microscope to qualitatively determine efficiency, or flow cytometry could be potentially used to sort and collect the GFP positive population for future experiments. It is speculated that cells that had successful knockdown of FAM83H-AS1 were the ones that underwent apoptosis, therefore there would be technical difficulties if this was true; if knockdown efficiency was increased then there would not be any viable cells to conduct experiments with. One alternative strategy could be the knockdown of FAM83H-AS1 using tetracycline-inducible shRNA viral vector that could express an siRNA against FAM83H-AS1 in a dose-dependent manner. If successful knockdown of FAM83H-AS1 induced cellular death in cells with knockdown, however, it would be a beneficial therapeutic approach to induce cell death in cancer cells. It would be critical to determine the effect of FAM83H-AS1 knockdown on surrounding normal and microenvironment cells, however, before moving forward. Simply transfecting normal cervical keratinocytes with siRNA against FAM83H-AS1 would be a good start to obtain this data.

If it is confirmed that knockdown of FAM83H-AS1 kills all the cells, it would eliminate the potential *in vivo* experimental set-up to knockdown siRNA in cervical cancer cells prior to injecting into a human tumor to see the reduction in tumor formation. It would be ideal to develop an *in vivo* system to develop a mouse tumor and then reduce FAM83H-AS1 expression and observe alterations in tumor growth. Currently, patient-derived xenograft (PDX) murine models are utilized in the cancer field as they allow for direct engraftment of human tumor tissues. A group previously showed the utilization of a PDX model with cervical cancer tumor engraftment(45). The Martinez lab has collaborated with West Virginia University

(WVU) PDX core to establish PDX cervical cancer tumors. It would be ideal to have the molecular background (FAM83H-AS1 expression) of the human tumors prior to implantation into the mouse so we could compare therapeutic effects with higher and lower FAM83H-AS1 expression in the tumor. Additionally, we could compare the therapeutic effects observed in our model to the human clinical outcomes. In the future, siRNA against FAM83H-AS1 could be injected into the mice and tumor size could be monitored. There is a collaborator at WVU with expertise in targeted drug delivery with nanoparticles (Dr. Werner Geldenhuys). Theoretically, siRNAs could be bound to nanoparticles to increase the internalization of siRNA into the nucleus. FAM83H-AS1 knockdown could be combined with current chemotherapies to improve their efficacy. If benefits are observed with FAM83H-AS1 knockdown with siRNAs, further studies could be developed on a drug to target FAM83H-AS1 that might be more clinically appealing than siRNAs.

Overall, FAM83H-AS1 regulates many of the cellular hallmarks of cancer in a variety of different cancers and should be studied further as a therapeutic target, diagnostic, and/or prognostic marker.

Another interesting phenomenon is that artificial inhibition of the HPV oncoproteins E6 and E7 in cervical cancer cells can induce cellular growth arrest by recovering the expression of the wild-type p53 and Rb tumor suppressor genes. More specifically, Martinez et al. showed that inhibition of HPV E7 can induce an irreversible form of growth arrest known as senescence through the Rb pathway by regulating specific miRNAs(46). Another important form of reversible growth arrest, known as quiescence, has been shown to be important in maintaining the homeostasis between differentiation and self-renewal of stem cells, and dysregulation of quiescence is related to cancer(47). Increasing knowledge on this cellular process will benefit research on aging-related diseases, wound healing treatment, organ regeneration, and cancer prevention. Our group, and others, have shown that ncRNAs (e.g. miRNAs and lncRNAs) are dysregulated in cellular quiescence (Chapter 3, Figure 1)(4).

Our group discovered that quiescent cells display an alternative miRNA biogenesis pathway in normal, primary foreskin fibroblast cells (Chapter 3), and further characterization of this pathway could reveal novel therapeutic and diagnostic markers. Currently, research on abnormal cellular growth arrest is focused on genetic alterations or mutations in cellular factors involved in the canonical pathway, and it has been demonstrated that mutations of some miRNA-processing factors have been linked to tumor formation(48). The findings in Chapter 3 show that there are alternations in the miRNA biogenesis pathway, which are not due to genetic mutations, that need to be researched further. It is hypothesized that quiescent cells disrupt the canonical miRNA biogenesis pathway to induce cellular growth arrest but continue the

production of essential miRNAs by the alternative pathway identified in Chapter 3. Other groups have demonstrated certain cells exhibit non-canonical miRNA processing(49, 50), but these non-canonical pathways only affect certain miRNAs without modifying the global canonical biogenesis pathway.

Specifically, our group discovered that Exportin-5 (XPO5) expression is reduced in quiescence by autophagy and miR-34a (Chapter 3, Figure 2), and certain miRNAs are induced in quiescence despite the absence of XPO5 (Chapter 3, Figure 3). Primary microRNAs (pri-miRNAs) corresponding to the quiescence-induced miRNAs are exported by Exportin-1 (XPO1) (Chapter 3, Figure 3) potentially via a trimethylguanosine ($m^{2,2,7}G$, TMG)-cap (Chapter 3, Figure 4).

Pri-miRNAs are canonically processed into precursor miRNAs (pre-miRNAs) in the nucleus(51), however certain pri-miRNAs (e.g. pri-miR-34a) were identified in the cytoplasm in quiescent cells (Chapter 3, Figure 5). The components involved in processing of these pri-miRNAs are not fully characterized. It is unclear if the pri-miRNA cap modification of these quiescence-induced miRNAs occurs in the nucleus or the cytoplasm. Future studies could identify the molecular components (and their localization) necessary for TMG capping of these pri-miRNA. (m^7G)-capped nuclear RNAs [small nuclear RNAs (snRNAs)] can be exported by XPO1 in collaboration with cap-binding complex (CBC) and the phosphorylated adaptor for RNA export (PHAX)(52). In the cytoplasm, these snRNAs are recognized by Sec1/Munc18 (Sm) proteins and the cap is hypermethylated into a ($m^{2,2,7}G$, TMG)-cap by TGS1. These snRNAs are then imported back into the nucleus by Snurportin-1 in association with other factors(53). It is hypothesized that quiescence-induced pri-miRNAs could be capped in a similar fashion. To test this hypothesis, cellular localization of ($m^{2,2,7}G$, TMG)-capped pri-miRNAs will be determined. Briefly, quiescent cells will be fractionated, an anti-TMG cap antibody will be used to perform RNA immunoprecipitation (RIP) and recovered RNA will be sent for RNA-seq. Our group published a protocol to immunoprecipitate TMG-capped RNA that could be used (Appendix). It will be determined if Sm protein is directly binding to these pri-miRNAs in the cytoplasmic fraction by performing RIP assays and mutational experiments.

Typically, Drosha does not cleave pri-miRNAs alone, as it forms a nuclear microprocessor complex with DGCR8 in the nucleus of cells to recognize and cleave pri-miRNAs and form pre-miRNAs. DGCR8 acts as a molecular anchor necessary for the recognition of pri-miRNAs at the double-stranded RNA junction and directs Drosha to cleave and release the hairpin-shape pre-miRNAs in the nucleus(54). Our group additionally showed that there is a cytoplasmic short form of Drosha in quiescent cells that is functional and potentially involved in the processing (Chapter 3, Figure 6), thus it was speculated that DGCR8 would also be detected in the cytoplasm. Interestingly, we were unable to detect DGCR8 in the

cytoplasm during quiescence (Chapter 3, Figure 6B). DGCR8 was knocked down by siRNA in quiescent cells and the biogenesis of quiescence-induced miR-34a was not affected (Chapter 3, Supplementary Figure 14B). It is possible, however, that DGCR8 in the cytoplasm has a cleaved form present in the cytoplasm (similar to Drosha) that was not targeted by the siRNA against DGCR8 or detected by the specific DGCR8 antibody used. If this is not the case, it is speculated that there are alternative DGCR8-like, double stranded RNA binding proteins in the cytoplasm aiding Drosha in miRNA processing. A future study to characterize the Drosha-cytoplasmic microprocessor complex during cellular quiescence would be to conduct co-immunoprecipitations using an antibody against Drosha in cytoplasmic fractions of quiescent cells, followed by mass spectroscopy. Once a DGCR8-like protein(s) is identified as an interactor with Drosha in the cytoplasm, it would need to be proven to be involved in the pri-miRNA processing. Briefly, the DGCR8-like protein(s) could be knocked out by CRISPR/Cas9 technology and downstream pri-miRNA and mature miRNA levels could be monitored.

A specific group of miRNAs is induced during cellular quiescence. Individually identifying genes targeted by these individual miRNAs could identify critical regulators in cellular growth arrest. It would be interesting to further identify global regulation by these induced miRNAs during quiescence. To identify interaction between the active miRNAs and their real biological mRNAs, high-throughput sequencing of RNA isolated by crosslinking immunoprecipitation (HITS-CLIP) could be conducted in quiescent cells.

Normal human foreskin fibroblasts were used in Chapter 3 to identify and study the non-canonical miRNA biogenesis pathway in quiescent cells. To broaden the applications, future studies analyzing miRNA biogenesis of stem cells and cancer stem cells, which undergo quiescence before they differentiate or start self-renewal, could be conducted.

Overall, further elucidation of this quiescence-induced miRNA biogenesis pathway could identify specific novel target genes that are involved in cellular proliferation, apoptosis inhibition, and autophagy pathways and could be potential therapeutic targets or diagnostic markers. These findings could thus be globally impactful, as essentially every cell in the human body has the potential to undergo these critical cellular processes.

References

1. **Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, Thurman RE, Kuehn MS, Taylor CM, Neph S, Koch CM, Asthana S, Malhotra A, Adzhubei I, Greenbaum JA, Andrews RM, Flicek P, Boyle PJ, Cao H, Carter NP, Clelland GK, Davis S, Day N, Dhami P, Dillon SC, Dorschner MO, Fiegler H, Giresi PG, Goldy J, Hawrylycz M, Haydock A, Humbert R, James KD, Johnson BE, Johnson EM, Frum TT, Rosenzweig ER, Karnani N, Lee K, Lefebvre GC, Navas PA, Neri F, Parker SC, Sabo PJ, Sandstrom R, Shafer A, Vetrie D, et al.** 2007. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* **447**:799-816.
2. **Wang H, Zhao Y, Chen M, Cui J.** 2017. Identification of Novel Long Non-coding and Circular RNAs in Human Papillomavirus-Mediated Cervical Cancer. *Front Microbiol* **8**:1720.
3. **Huang J, Liu T, Shang C, Zhao Y, Wang W, Liang Y, Guo L, Yao S.** 2018. Identification of lncRNAs by microarray analysis reveals the potential role of lncRNAs in cervical cancer pathogenesis. *Oncol Lett* **15**:5584-5592.
4. **Bierhoff H, Dammert MA, Brocks D, Dambacher S, Schotta G, Grummt I.** 2014. Quiescence-induced lncRNAs trigger H4K20 trimethylation and transcriptional silencing. *Mol Cell* **54**:675-682.
5. **Berman TA, Schiller JT.** 2017. Human papillomavirus in cervical cancer and oropharyngeal cancer: One cause, two diseases. *Cancer* **123**:2219-2229.
6. **de Martel C, Plummer M, Vignat J, Franceschi S.** 2017. Worldwide burden of cancer attributable to HPV by site, country and HPV type. *Int J Cancer* **141**:664-670.
7. **Schiller JT, Muller M.** 2015. Next generation prophylactic human papillomavirus vaccines. *Lancet Oncol* **16**:e217-225.
8. **Bruni L, Diaz M, Barrionuevo-Rosas L, Herrero R, Bray F, Bosch FX, de Sanjose S, Castellsague X.** 2016. Global estimates of human papillomavirus vaccination coverage by region and income level: a pooled analysis. *Lancet Glob Health* **4**:e453-463.
9. **Munger K, Howley PM.** 2002. Human papillomavirus immortalization and transformation functions. *Virus Res* **89**:213-228.
10. **Moody CA, Laimins LA.** 2010. Human papillomavirus oncoproteins: pathways to transformation. *Nat Rev Cancer* **10**:550-560.
11. **Scheffner M, Werness BA, Huibregtse JM, Levine AJ, Howley PM.** 1990. The E6 oncoprotein encoded by human papillomavirus types 16 and 18 promotes the degradation of p53. *Cell* **63**:1129-1136.
12. **Boyer SN, Wazer DE, Band V.** 1996. E7 protein of human papilloma virus-16 induces degradation of retinoblastoma protein through the ubiquitin-proteasome pathway. *Cancer Res* **56**:4620-4624.
13. **Yim EK, Park JS.** 2005. The role of HPV E6 and E7 oncoproteins in HPV-associated cervical carcinogenesis. *Cancer Res Treat* **37**:319-324.
14. **Martinez I, Gardiner AS, Board KF, Monzon FA, Edwards RP, Khan SA.** 2008. Human papillomavirus type 16 reduces the expression of microRNA-218 in cervical carcinoma cells. *Oncogene* **27**:2575-2582.
15. **Jiang Y, Li Y, Fang S, Jiang B, Qin C, Xie P, Zhou G, Li G.** 2014. The role of MALAT1 correlates with HPV in cervical cancer. *Oncol Lett* **7**:2135-2141.
16. **Sharma S, Munger K.** 2018. Expression of the cervical carcinoma expressed PCNA regulatory (CCEPR) long noncoding RNA is driven by the human papillomavirus E6 protein and modulates cell proliferation independent of PCNA. *Virology* **518**:8-13.
17. **White EA, Kramer RE, Tan MJ, Hayes SD, Harper JW, Howley PM.** 2012. Comprehensive analysis of host cellular interactions with human papillomavirus E6 proteins identifies new E6 binding partners and reflects viral diversity. *J Virol* **86**:13174-13186.

18. **Masson M, Hindelang C, Sibler AP, Schwalbach G, Trave G, Weiss E.** 2003. Preferential nuclear localization of the human papillomavirus type 16 E6 oncoprotein in cervical carcinoma cells. *J Gen Virol* **84**:2099-2104.
19. **Yan X, Hu Z, Feng Y, Hu X, Yuan J, Zhao SD, Zhang Y, Yang L, Shan W, He Q, Fan L, Kandalaft LE, Tanyi JL, Li C, Yuan CX, Zhang D, Yuan H, Hua K, Lu Y, Katsaros D, Huang Q, Montone K, Fan Y, Coukos G, Boyd J, Sood AK, Rebbeck T, Mills GB, Dang CV, Zhang L.** 2015. Comprehensive Genomic Characterization of Long Non-coding RNAs across Human Cancers. *Cancer Cell* **28**:529-540.
20. **Cabanski CR, White NM, Dang HX, Silva-Fisher JM, Rauck CE, Cicka D, Maher CA.** 2015. Pan-cancer transcriptome analysis reveals long noncoding RNAs with conserved function. *RNA Biol* **12**:628-642.
21. **Yang F, Lv SX, Lv L, Liu YH, Dong SY, Yao ZH, Dai XX, Zhang XH, Wang OC.** 2016. Identification of lncRNA FAM83H-AS1 as a novel prognostic marker in luminal subtype breast cancer. *Onco Targets Ther* **9**:7039-7045.
22. **Lu S, Dong W, Zhao P, Liu Z.** 2018. lncRNA FAM83H-AS1 is associated with the prognosis of colorectal carcinoma and promotes cell proliferation by targeting the Notch signaling pathway. *Oncol Lett* **15**:1861-1868.
23. **Yang L, Xu L, Wang Q, Wang M, An G.** 2016. Dysregulation of long non-coding RNA profiles in human colorectal cancer and its association with overall survival. *Oncol Lett* **12**:4068-4074.
24. **Zhang J, Feng S, Su W, Bai S, Xiao L, Wang L, Thomas DG, Lin J, Reddy RM, Carrott PW, Lynch WR, Chang AC, Beer DG, Guo YM, Chen G.** 2017. Overexpression of FAM83H-AS1 indicates poor patient survival and knockdown impairs cell proliferation and invasion via MET/EGFR signaling in lung cancer. *Sci Rep* **7**:42819.
25. **Arnes L, Liu Z, Wang J, Carlo Maurer H, Sagalovskiy I, Sanchez-Martin M, Bommakanti N, Garofalo DC, Balderes DA, Sussel L, Olive KP, Rabadan R.** 2018. Comprehensive characterisation of compartment-specific long non-coding RNAs associated with pancreatic ductal adenocarcinoma. *Gut* doi:10.1136/gutjnl-2017-314353.
26. **Bi YY, Shen G, Quan Y, Jiang W, Xu F.** 2018. Long noncoding RNA FAM83H-AS1 exerts an oncogenic role in glioma through epigenetically silencing CDKN1A (p21). *J Cell Physiol* doi:10.1002/jcp.26813.
27. **Guan P, Howell-Jones R, Li N, Bruni L, de Sanjose S, Franceschi S, Clifford GM.** 2012. Human papillomavirus types in 115,789 HPV-positive women: a meta-analysis from cervical infection to cancer. *Int J Cancer* **131**:2349-2359.
28. **Graham SV, Faizo AAA.** 2017. Control of human papillomavirus gene expression by alternative splicing. *Virus Res* **231**:83-95.
29. **Biegging KT, Mello SS, Attardi LD.** 2014. Unravelling mechanisms of p53-mediated tumour suppression. *Nat Rev Cancer* **14**:359-370.
30. **Patel D, Huang SM, Baglia LA, McCance DJ.** 1999. The E6 protein of human papillomavirus type 16 binds to and inhibits co-activation by CBP and p300. *Embo j* **18**:5061-5072.
31. **Zimmermann H, Degenkolbe R, Bernard HU, O'Connor MJ.** 1999. The human papillomavirus type 16 E6 oncoprotein can down-regulate p53 activity by targeting the transcriptional coactivator CBP/p300. *J Virol* **73**:6209-6219.
32. **Kreft L, Soete A, Hulpiau P, Botzki A, Saeys Y, De Bleser P.** 2017. ConTra v3: a tool to identify transcription factor binding sites across species, update 2017. *Nucleic Acids Res* **45**:W490-w494.
33. **Fermento ME, Gandini NA, Salomon DG, Ferronato MJ, Vitale CA, Arevalo J, Lopez Romero A, Nunez M, Jung M, Facchinetti MM, Curino AC.** 2014. Inhibition of p300 suppresses growth of breast cancer. Role of p300 subcellular localization. *Exp Mol Pathol* **97**:411-424.
34. **Ishihama K, Yamakawa M, Semba S, Takeda H, Kawata S, Kimura S, Kimura W.** 2007. Expression of HDAC1 and CBP/p300 in human colorectal carcinomas. *J Clin Pathol* **60**:1205-1210.

35. **Hou X, Li Y, Luo RZ, Fu JH, He JH, Zhang LJ, Yang HX.** 2012. High expression of the transcriptional co-activator p300 predicts poor survival in resectable non-small cell lung cancers. *Eur J Surg Oncol* **38**:523-530.
36. **Ono H, Basson MD, Ito H.** 2016. P300 inhibition enhances gemcitabine-induced apoptosis of pancreatic cancer. *Oncotarget* **7**:51301-51310.
37. **Panicker SP, Raychaudhuri B, Sharma P, Tipps R, Mazumdar T, Mal AK, Palomo JM, Vogelbaum MA, Haque SJ.** 2010. p300- and Myc-mediated regulation of glioblastoma multiforme cell differentiation. *Oncotarget* **1**:289-303.
38. **Sigova AA, Mullen AC, Molinie B, Gupta S, Orlando DA, Guenther MG, Almada AE, Lin C, Sharp PA, Giallourakis CC, Young RA.** 2013. Divergent transcription of long noncoding RNA/mRNA gene pairs in embryonic stem cells. *Proc Natl Acad Sci U S A* **110**:2876-2881.
39. **Khachane AN, Harrison PM.** 2010. Mining mammalian transcript data for functional long non-coding RNAs. *PLoS One* **5**:e10316.
40. **Snijders AM, Lee SY, Hang B, Hao W, Bissell MJ, Mao JH.** 2017. FAM83 family oncogenes are broadly involved in human cancers: an integrative multi-omics approach. *Mol Oncol* **11**:167-179.
41. **Yu B, Shan G.** 2016. Functions of long noncoding RNAs in the nucleus. *Nucleus* **7**:155-166.
42. **Chen G, Peng L, Zhu Z, Du C, Shen Z, Zang R, Su Y, Xia Y, Tang W.** 2017. LncRNA AFAP1-AS Functions as a Competing Endogenous RNA to Regulate RAP1B Expression by sponging miR-181a in the HSCR. *Int J Med Sci* **14**:1022-1030.
43. **Liu P, Yang H, Zhang J, Peng X, Lu Z, Tong W, Chen J.** 2017. The lncRNA MALAT1 acts as a competing endogenous RNA to regulate KRAS expression by sponging miR-217 in pancreatic ductal adenocarcinoma. *Sci Rep* **7**:5186.
44. **Konermann S, Brigham MD, Trevino AE, Joung J, Abudayyeh OO, Barcena C, Hsu PD, Habib N, Gootenberg JS, Nishimasu H, Nureki O, Zhang F.** 2015. Genome-scale transcriptional activation by an engineered CRISPR-Cas9 complex. *Nature* **517**:583-588.
45. **Chaudary N, Pintilie M, Schwock J, Dhani N, Clarke B, Milosevic M, Fyles A, Hill RP.** 2012. Characterization of the Tumor-Microenvironment in Patient-Derived Cervix Xenografts (OCICx). *Cancers (Basel)* **4**:821-845.
46. **Martinez I, Cazalla D, Almstead LL, Steitz JA, DiMaio D.** 2011. miR-29 and miR-30 regulate B-Myb expression during cellular senescence. *Proc Natl Acad Sci U S A* **108**:522-527.
47. **Cheung TH, Rando TA.** 2013. Molecular regulation of stem cell quiescence. *Nat Rev Mol Cell Biol* **14**:329-340.
48. **Melo SA, Moutinho C, Ropero S, Calin GA, Rossi S, Spizzo R, Fernandez AF, Davalos V, Villanueva A, Montoya G, Yamamoto H, Schwartz S, Jr., Esteller M.** 2010. A genetic defect in exportin-5 traps precursor microRNAs in the nucleus of cancer cells. *Cancer Cell* **18**:303-315.
49. **Havens MA, Reich AA, Duelli DM, Hastings ML.** 2012. Biogenesis of mammalian microRNAs by a non-canonical processing pathway. *Nucleic Acids Res* **40**:4626-4640.
50. **Shapiro JS, Langlois RA, Pham AM, Tenoever BR.** 2012. Evidence for a cytoplasmic microprocessor of pri-miRNAs. *Rna* **18**:1338-1346.
51. **Lee Y, Ahn C, Han J, Choi H, Kim J, Yim J, Lee J, Provost P, Radmark O, Kim S, Kim VN.** 2003. The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**:415-419.
52. **Boulon S, Verheggen C, Jady BE, Girard C, Pescia C, Paul C, Ospina JK, Kiss T, Matera AG, Bordonne R, Bertrand E.** 2004. PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli. *Mol Cell* **16**:777-787.
53. **Kiss T.** 2004. Biogenesis of small nuclear RNPs. *J Cell Sci* **117**:5949-5951.
54. **Gregory RI, Yan KP, Amuthan G, Chendrimada T, Doratotaj B, Cooch N, Shiekhattar R.** 2004. The Microprocessor complex mediates the genesis of microRNAs. *Nature* **432**:235-240.

Appendix

Immunoprecipitation of Tri-methylated Capped RNA

Karen E. Hayes¹, Jamie A. Barr¹, Mingyi Xie², Joan A. Steitz^{3, 4, 5} and Ivan Martinez^{1, *}

¹Department of Microbiology, Program in Cancer Cell Biology, WVU Cancer Institute, West Virginia University, Morgantown, WV, USA; ²Department of Biochemistry & Molecular Biology, U.F. Health Cancer Center, University of Florida, Gainesville, FL, USA; ³Department of Molecular Biophysics & Biochemistry, Yale University, New Haven, CT, USA; ⁴Howard Hughes Medical Institute, Yale University, New Haven, CT 06536, USA; ⁵Yale Cancer Center, New Haven, CT, USA

*For correspondence: ivmartinez@hsc.wvu.edu

Keywords: m^{2,2,7}G-cap RNA, TMG-cap RNA, Tri-methylated RNA, RNA immunoprecipitation, Pri-miRNA

Published in: Bio-protocol journal
Bio Protoc. 2018 Feb 5;8(3). PMID: 29527542

I. Abstract

Cellular quiescence (also known as G₀ arrest) is characterized by reduced DNA replication, increased autophagy, and increased expression of cyclin-dependent kinase p27^{Kip1}. Quiescence is essential for wound healing, organ regeneration, and preventing neoplasia. Previous findings indicate that microRNAs (miRNAs) play an important role in regulating cellular quiescence. Our recent publication demonstrated the existence of an alternative miRNA biogenesis pathway in primary human foreskin fibroblast (HFF) cells during quiescence. Indeed, we have identified a group of pri-miRNAs (whose mature miRNAs were found induced during quiescence) modified with a 2,2,7-trimethylguanosine (TMG)-cap by the trimethylguanosine synthase 1 (TGS1) protein and transported to the cytoplasm by the Exportin-1 (XPO1) protein. We used an antibody against (TMG)-caps (which does not cross-react with the (m⁷G)-caps that most pri-miRNAs or mRNAs contain [Luhrmann *et al.*, 1982]) to perform RNA immunoprecipitations from total RNA extracts of proliferating or quiescent HFFs. The novelty of this assay is the specific isolation of pri-miRNAs as well as other non-coding RNAs containing a TMG-cap modification.

II. Background

Cellular quiescence, a type of reversible growth arrest, is an important cellular state involved in wound healing, organ regeneration, and preventing neoplasia (Coller, 2011; Valcourt *et al.*, 2012). Small non-coding RNAs such as miRNAs have been found involved in the regulation of cellular quiescence. miRNAs are small non-coding RNAs ~22-nucleotides long that regulate the expression of protein-coding genes by base-pairing with the 3' untranslated region (3'UTR) of messenger RNAs (mRNAs) (Esteller, 2011). The canonical miRNA biogenesis pathway is based on a stepwise processing machinery (Ha and Kim, 2014; Kim *et al.*, 2016). miRNAs are transcribed to produce a primary miRNA (pri-miRNA) with an imperfect loop structure that is recognized by the enzyme Drosha and its binding partner DGCR8 in the nucleus. Cleavage of the pri-miRNA generates a precursor miRNA (pre-miRNA) that is recognized and transported to the cytoplasm by the Exportin-5 (XPO5) protein. The pre-miRNA is cleaved by the enzyme Dicer (mature miRNA) and loaded into the RNA-induced silencing complex (RISC). On the other hand, precursors of small nuclear RNAs (snRNAs) involved in mRNA processing such as U1, U2, U4, and U5 have a (m⁷G)-cap, which is recognized by cap-binding complex (CBC) and the phosphorylated adaptor for RNA export (PHAX) in the nucleus to enable their export to the cytoplasm by XPO1 (Ohno *et al.*, 2000). These snRNAs are then recognized by Sec1/Munc18 (Sm) proteins (by binding to Sm binding site sequences) in the cytoplasm and TGS1 is recruited to hypermethylate the (m⁷G)-cap into a (m^{2,2,7}G, TMG)-cap. This modification is recognized by Snuportin-1 in association with Importin-β and other factors to

import the snRNAs back into the nucleus (Palacios *et al.*, 1997; Kiss, 2004). Interestingly, XPO1 also has high affinity for the (TMG)-capped small nucleolar RNA (snoRNA) U3 in the nucleus and transports it from Cajal bodies to the nucleoli (Boulon *et al.*, 2004). A previous study showed that TGS1 enhances Rev-dependent HIV-1 RNA expression by (TMG)-capping viral mRNAs in the nucleus, thereby increasing recognition by XPO1 for transport to the cytoplasm (Yedavalli and Jeang, 2010). These findings suggest that TMG-capping of RNAs gives plasticity to different types of RNA molecules in order to regulate their processing and cellular localization. Our recent findings demonstrated the existence of a group of pri-miRNAs modified with a 2,2,7-trimethylguanosine (TMG)-cap by TGS1 protein and transported to the cytoplasm by XPO1 during quiescence. Previous publications have shown the ability to pull-down (TMG)-cap RNAs, such as snRNAs and snoRNAs, with specific antibodies against (TMG)-cap RNAs (Luhrmann *et al.*, 1982). Our previous publications demonstrated for the first time the pull-down of (TMG)-cap pri-miRNAs in human cells (Martinez *et al.*, 2017). Understanding which RNAs could be modified with a TMG-cap will provide new important insights into RNA biogenesis in normal or disease-related conditions.

C. Materials and Reagents

1. 1.5 ml microcentrifuge tubes (Fisher Scientific, catalog number: [05-408-129](#))
2. 15 ml conical centrifuge tubes (DNase-/RNase-free) (Corning, catalog number: [430052](#))
3. Gilson™ EXPERT™ University Fit pipette filter tips (Gilson, catalog numbers: [F1731031](#), [F1733031](#), [F1735031](#), [F1737031](#))
4. Gel-loading pipet tips (Fisher Scientific, catalog number: [02-707-139](#))
5. Large-orifice pipet tips (Fisher Scientific, catalog number: [02-707-134](#))
6. Sterile polystyrene disposable serological pipettes (Greiner Bio One International, catalog number: [710180](#))
7. Serological pipettes
 - 2 ml serological pipettes (Fisher Scientific, catalog number: [13-678-11C](#))
 - 5 ml serological pipettes (Fisher Scientific, catalog number: [13-678-11D](#))
 - 10 ml serological pipettes (Fisher Scientific, catalog number: [13-678-11E](#))
 - 25 ml serological pipettes (Fisher Scientific, catalog number: [13-678-11](#))
8. 150 cm² vented tissue culture treated flasks (Corning, Falcon®, catalog number: [355001](#))
9. 100 mm TC-treated cell culture dish (Corning, Falcon®, catalog number: [353033](#))
10. Cell Scrapers (Fisher Scientific, catalog number: [08-100-242](#))
11. HFF cells (obtained from the Yale Skin Disease Research Center) (Alternative source of HFF cells from ATCC: Hs27 (ATCC, catalog number: [CRL-1634](#)))

12. DMEM (Sigma-Aldrich, catalog number: [D7777](#))
13. Minimum essential medium (MEM) non-essential amino acids, 100x (Thermo Fisher Scientific, catalog number: [11140076](#))
14. 0.05% Trypsin-EDTA with phenol red (Thermo Fisher Scientific, catalog number: [25300054](#))
15. 10x PBS (Sigma-Aldrich, catalog number: [P5493](#))
16. Nuclease-free water (not DEPC-Treated) (Thermo Fisher Scientific, catalog number: [AM9937](#))
17. TRIzol™ Reagent (Thermo Fisher Scientific, catalog number: [15596026](#))
18. RNase AWAY™ Surface Decontaminant (Thermo Fisher Scientific, catalog number: [7002](#))
19. Chloroform (Sigma-Aldrich, catalog number: [C2432](#))
20. Ethanol, molecular biology grade (Fisher Scientific, catalog number: [BP2818-500](#))
21. Isopropanol, molecular biology grade (Fisher Scientific, catalog number: [BP26184](#))
22. GlycoBlue™ Coprecipitant (15 mg/ml) (Thermo Fisher Scientific, catalog number: [AM9516](#))
23. TURBO DNA-free™ Kit (Thermo Fisher Scientific, catalog number: [AM1907](#))
24. Anti-m3G-cap, rabbit polyclonal, antiserum (Synaptic Systems)*

*Note: *Synaptic Systems discontinued the production of the Anti-m3G-cap, rabbit polyclonal, antiserum. Creative Diagnostics has a rabbit anti-TMG antibody (Anti-m3G-cap polyclonal antibody, Creative Diagnostics, catalog number: [DPAB29202](#)) that would be similar to the one we previously used but the experimental conditions have to be re-evaluated.*

As we mentioned above, the company (Synaptic Systems) discontinued this antibody. For that reason, we recommend to use a similar antibody from Creative Diagnostics. We agree to not mention the catalog number of the previous company to avoid confusion.

25. Protein G Sepharose® 4 Fast Flow Beads (GE Healthcare, catalog number: [17061801](#))
26. Normal rabbit serum (control, EMD Millipore, catalog number: [NS01L-1ML](#))
27. NaCl (Fisher Scientific, catalog number: [S671-3](#))
28. NP-40 (Thermo Fisher Scientific, catalog number: [28324](#))
29. Tris base (Fisher Scientific, catalog number: [BP152-5](#))
30. HCl (VWR, catalog number: [BDH7204-1](#))
31. RNasin™ Plus RNase inhibitor (Promega, catalog number: [N2611](#))
32. NaOAc (Amresco, catalog number: 0602)
33. EDTA, pH 8 (Thermo Fisher Scientific, catalog number: [AM9260G](#))
34. EDTA (Amresco, catalog number: 0105)
35. SDS (Thermo Fisher Scientific, catalog number: [AM9822](#))
36. Phenol/Chloroform/Isoamyl Alcohol; 125:24:1 mixture, pH 4.5 (Thermo Fisher Scientific, catalog number: [AM9720](#))

37. Agarose LE (Denville Scientific, catalog number: [CA3510-8](#))
38. iScript™ cDNA Synthesis Kit (Bio-Rad Laboratories, catalog number: [1708891](#))
39. Sso Advanced™ Universal SYBR® Green Supermix (Bio-Rad Laboratories, catalog number: [1725274](#))
40. PARIS™ Kit (Thermo Fisher Scientific, catalog number: [AM1921](#))
41. Boric acid (Fisher Scientific, catalog number: [A73-1](#))
42. NET-2 Buffer (see Recipes)
43. G-50 Buffer (see Recipes)
44. 1x TBE (see Recipes)

D. Equipment

1. Micropipettes (Gilson, model: Pipetman® L, catalog number: F167370)
2. Forma™ Steri-Cycle™ CO₂ Incubator (Thermo Scientific, model: [Forma™ Steri-Cycle™ CO₂ Incubators](#), catalog number: 370)
3. -80 °C freeze
4. Sorvall™ Legend™ Micro 21R Microcentrifuge (Thermo Fisher Scientific, model: [Sorvall™ Legend™ Micro 21R](#), catalog number: 75002490)
5. Eppendorf™ Thermomixer™ R (Eppendorf, model: Thermomixer R, catalog number: 05-412-401)
6. Labquake™ Tube Shaker/Rotator (Thermo Fisher Scientific, catalog number: C4152110Q)
7. Sorvall™ ST 40R Centrifuge (Thermo Fisher Scientific, model: Sorvall™ ST 40R, catalog number: 75004525)
8. NanoDrop™ 2000 Spectrophotometer (Thermo Fisher Scientific, model: [NanoDrop™ 2000](#), catalog number: ND-2000)
9. T100™ Thermal Cycler (Bio-Rad Laboratories, catalog number: [1861096](#))
10. CFX Connect™ Real-Time PCR Detection System (Bio-Rad Laboratories, catalog number: [1855200](#))
11. UV transilluminator

E. Procedure

1. Plate HFF cells in serum-free DMEM supplemented with MEM non-essential amino acids at low density (approx. 1×10^6 cells/150 cm² flask). Incubate HFF cells in humidified atmosphere of 5% CO₂ at 37 °C for 48 h (Figure 1).

Notes:

- a. *In order to obtain enough RNA for each experiment, use 4 to 6 150 cm² flasks.*
 - b. *To plate HFF cells in serum-free DMEM; wash 70-80% confluent HFF cells (100 mm TC-treated cell culture dish) three times each with 10 ml phosphate buffered saline, add 1 ml 0.05% trypsin with phenol red to detach cells, add 10 ml serum-free DMEM media and collect cells in 15 ml conical tube, pellet cells by centrifuging at 300 x g at room temperature, and re-suspend pellet in serum-free DMEM supplemented medium.*
 - c. *It is important to remove trypsin from HFF cells before re-plating since medium doesn't contain FBS to inactive the trypsin.*
2. Collect HFF cells by scraping flask on ice, pellet at 300 x g for 5 min at 4 °C, and wash two times each with 2 ml of phosphate buffered saline. Cell pellets can be frozen at -80 °C until RNA extraction (cell pellets must be kept on ice until lysed with TRIzol™) (Figure 1).

Notes: To pellet HFF cells, remove old serum-free DMEM from cells and add 3 ml of fresh serum-free DMEM. Use the blade of cell scraper to collect cells in one corner of flask taking care to not splash the cells in flask. Add 1-2% of volume of FBS or BSA to cells on ice to assist in pelleting. Extract RNA from pelleted cells using TRIzol™ Reagent per manufacturer's instructions:

Note: All steps in this protocol should be performed in an RNase free bench area by using RNase AWAY™ Surface Decontaminant to spray and wipe most surface areas including micropipettes and ice bucket.

- a. Resuspend cell pellets thawed on ice in 1 ml of TRIzol™ Reagent for 5 min to lyse.
- b. Add chloroform (20% of TRIzol™ Reagent volume), mix samples well, and incubate at room temperature for 2-3 min.
- c. Centrifuge samples at 12,000 x g for 15 min at 4 °C to separate red phenol-chloroform (red/lower layer), interphase, and aqueous phase (clear/upper layer).
- d. Collect aqueous layer and precipitate RNA for 10 min at room temperature with isopropanol (0.5 volume of original TRIzol™ Reagent); GlycoBlue™ Coprecipitant (15 µg) can be added at this time to aid RNA precipitation.
- e. Pellet precipitated RNA by centrifugation (12,000 x g for 10 min at 4 °C).
- f. Wash RNA pellets with 75% ethanol (add an equal volume of 75% ethanol to TRIzol™ Reagent and briefly vortex) and re-pellet by centrifuging at 7,500 x g for 5 min at 4 °C.
- g. Air-dry RNA pellets for a few minutes and re-suspend in nuclease-free water.

Note: Do not completely dry the RNA pellet because it will be difficult to re-suspend. Incubation at 55 °C for 10 min can help solubilize the RNA.

3. Use DNase from TURBO DNA-free™ Kit to digest DNA contamination in RNA samples according to manufacturer's protocol:
 - a. Mix RNA with 10x TURBO DNase Buffer (0.1 volume) and 1 µl of TURBO DNase.
 - b. Incubate RNA at 37 °C for 20 min and halt the reaction by adding DNase Inactivation Reagent (0.1 volume).
 - c. Incubate RNA samples for 5 min with intermittent mixing by flicking tube then centrifuge at 10,000 \times g at room temperature for 2 min.
 - d. Collect RNA (supernatant) and quantitate with Nanodrop™ 2000 Spectrophotometer.

Note: The average amount of total RNA obtained from each flask is 1.5-2 µg.

4. Pre-load Protein G Sepharose® 4 Fast beads with rabbit serum control or anti-m3G-cap antibody (Figure 1).

Note: Use large-orifice pipet tips or clip the tip off of a pipet tip to aid the transfer of Protein G Sepharose® 4 Fast beads.

- a. Wash 4 times the Protein G Sepharose® 4 Fast beads (40 µl slurry/IP sample) each with 1 ml NET-2 Buffer (see Recipes) using short 30 sec spins at 1000 \times g to pellet beads.

Note: Non-filtered gel-loading pipet tips can be used during 'wash' steps to help prevent accidental loss of beads.

- b. Dilute washed Protein G Sepharose beads in 500 µl of NET-2 Buffer and add rabbit anti-m3G-cap antibody (15 µl/IP sample, approximately 150 µg) or rabbit serum control (15 µl/IP sample, approximately 150 µg). To achieve equal loading of beads with antibodies only use one microcentrifuge tube/antibody; aliquot the beads into separate tubes after they are pre-loaded.
 - c. Incubate beads/antibody for 1.5 h at room temperature on tube rotator; make sure samples are actually mixing.
 - d. Remove excess antibody by washing Protein G Sepharose Beads 5 times each with 1 ml NET-2 Buffer (short 30 sec spins at 1,000 \times g); Resuspend beads in NET-2 Buffer (50 µl/sample).
 - e. Aliquot antibody-bound Protein G Sepharose Beads into 1.5 microcentrifuge tubes; use enough tubes to have one of each antibody (Anti-m3G-cap and control) per sample.

5. Pre-clear RNA with Protein G Sepharose Beads to reduce non-specific binding:

- a. Add RNasin Plus RNase inhibitor (1 µl) to RNA (10 µg/500 µl diluted with NET-2 buffer) and incubate for 3 to 5 min at 85 °C.

Note: Plunge RNA into ice immediately after heating to avoid refolding of RNA.

- b. After heating, add an additional 1 µl of RNasin Plus RNase inhibitor to RNA.

Note: Remember to wipe tubes with RNase AWAY™ Surface Decontaminant before opening to prevent possible contamination and degradation of RNA.

- c. Rotate RNA (10 µg/sample for each antibody) slowly (without vibration) with 40 µl of NET 2-washed Protein G Sepharose 4 Fast Flow beads (without antibody) for 2 h at 4 °C.
- d. Centrifuge RNA at 1,000 \times g for 2 min to remove pre-clearing beads.
6. Rotate pre-cleared RNA slowly (without vibration) with anti-m3G-cap or control pre-loaded Protein G Sepharose® 4 Fast beads for 4 to 16 h at 4 °C. Confirm that samples are mixing (Figure 1).
7. Collect the beads by centrifuging at 1,000 \times g for 2 min and remove supernatant (save 250 µl for RNA extraction).
8. Wash beads 5 times with 1 ml of NET-2 Buffer using short 30 sec spins at 1,000 \times g.
Note: RNasin Plus RNase inhibitor can be added to NET-2 Buffer (see Recipes) to prevent degradation of RNA.
9. Resuspend beads in 250 µl of G-50 Buffer (see Recipes) to elute TMG-capped RNA (Figure 1).
10. Purify RNA by Phenol-chloroform-isoamyl alcohol (PCI) extraction (also extract RNA from TMG-capped depleted supernatant)
 - a. Add 250 µl of Phenol-chloroform-isoamyl alcohol to RNA in G-50 Buffer.
 - b. Vortex RNA samples for 20 sec (10 up and 10 angled) and centrifuge samples for 10 min at 12,000 \times g at room temperature.
 - c. Transfer the aqueous phase with RNA (upper layer) to a clean microcentrifuge tube containing 50 µl of 3 M sodium acetate (NaOAc), pH 5.2.
 - d. Add 1 ml of 100% ethanol (precooled to -20 °C) to precipitate RNA; 1 µl of GlycoBlue can be added to help precipitate RNA and visualize pellets.
 - e. Incubate RNA at -20 °C for 48 h.
 - f. Pellet RNA by centrifuging sample at 12,000 \times g for 10 min at 4 °C.
 - g. Aspirate ethanol and wash RNA pellet with 500 µl of 75% ethanol (precooled to -20 °C).
 - h. Re-pellet purified RNA by short spin for 1 min at 12,000 \times g at 4 °C and aspirate 75% ethanol.
 - i. Dissolve the dried m3G-capped RNA pellet in nuclease-free water (20 to 30 µl).

Note: The average amount of pull-down m3G-capped RNA obtained from each experiment (5 flasks) is 1.5-2 µg.

11. RT-PCR of m3G-capped RNAs by using iScript™ cDNA Synthesis Kit and 2x SsoAdvanced™ SYBR® Green Supermix PCR according to manufacturer's protocols:
Note: Perform RT-PCR using un-processed RNA (input) and RNA extracted from supernatant to determine expression levels of RNA of interest (m3G-capped and uncapped).
 - a. Mix 5 µl of m3G-capped RNA with 5x iScript Reaction Mix (0.2 volume) and 1 µl of iScript Reverse Transcriptase.

- b. Process RNA to cDNA in a thermocycler by undergoing the following conditions: 25 °C for 5 min, 46 °C for 20 min, 95 °C for 1 min, and held at 4 °C. The reaction can be stored at 4 °C for short-term storage or -20 °C for long-term storage.
- c. Mix the resulting cDNA (around ¼ of the total cDNA reaction mix) with 2x SsoAdvanced™ SYBR® Green Supermix (0.5 volume) and the desired forward and reverse primers (12.5 µM stock; 0.04 volume).
- d. cDNA then undergoes the following conditions to produce a PCR product: 95 °C for 1.5 min; 40 cycles: 95 °C for 30 sec, 60 °C for 30 sec.
- e. The PCR product is ran on a 1% agarose gel (made with 1x TBE [see Recipes] and ethidium bromide) at 100 V for 2 h, and bands are imaged with a UV transilluminator.

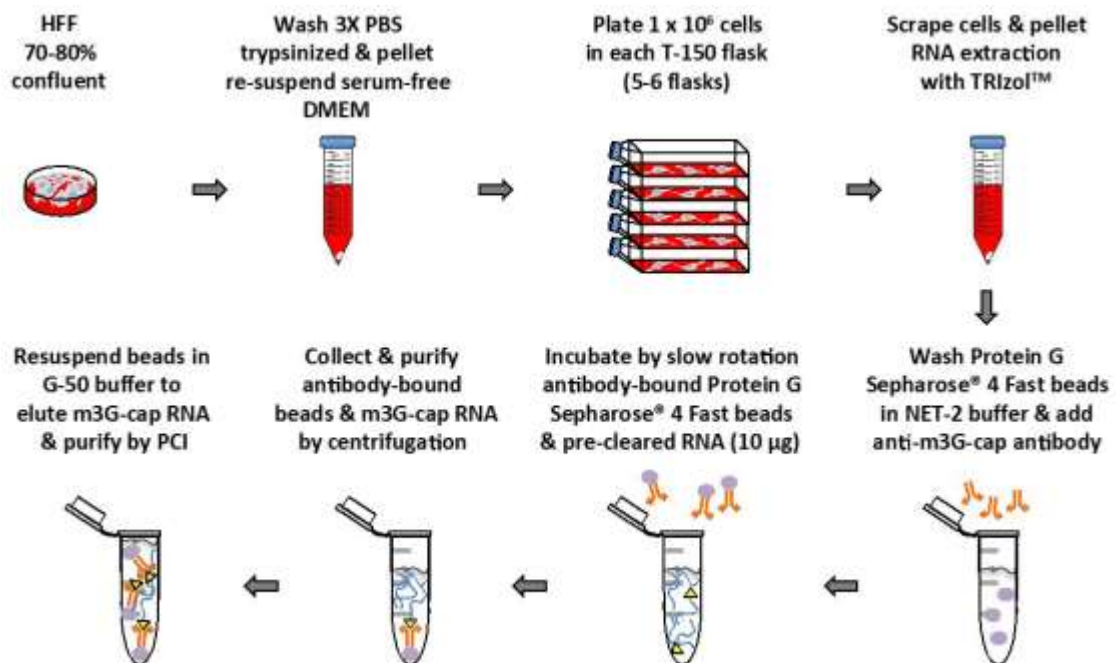


Figure 1. Main steps diagram of this protocol

F. Data analysis

The recovery of (TMG)-capped RNAs was measured by RT-PCR amplification of well-known hypermethylated RNAs such as small nuclear RNA (snRNA) U7 (Figure 2) or small nucleolar RNA (snoRNA) U3 (Martinez *et al.*, 2017).



Figure 2. RNA immunoprecipitation of (TMG)-capped primary miRNAs (Pri-miRNAs) in quiescent human foreskin fibroblasts (HFFs). RT-PCR data shows the RNA immunoprecipitation of Pri-miR-34a and Pri-miR-3188 in quiescent HFFs (as well as the positive control snRNA U7), but not Pri-miR-423 using an antibody against (TMG)-capped RNAs. Total RNA was extracted using TRIzol Reagent, and 10 µg of RNA was diluted in NET-2 buffer, precleared and incubated with Protein G Sepharose 4 Fast Flow beads loaded with 15 µl of control antibody (Normal Rabbit Serum, EMD-Millipore) or with antibody recognizing the (TMG)-cap (Anti-m3G-cap, rabbit polyclonal, Synaptic Systems). Beads were rinsed five times with NET-2 buffer and were resuspended in G-50 buffer. RNA was extracted from the beads by phenol-chloroform-isoamyl alcohol extraction and resuspended in 20 µl of nuclease-free water. Immunoprecipitated tri-methylated capped RNA was converted to cDNA using iScript cDNA synthesis kit (Bio-Rad), followed by RT-PCR, and visualized after gel electrophoresis.

G. Notes

This protocol could be modified to determine the location of TMG-capped RNA by separation of nuclear and cytoplasmic RNA fractions from fresh cultured cells (PARIS™ Kit).

H. Recipes

Note: All reagents should be made with nuclease-free water and autoclaved.

1. NET-2 Buffer
150 mM NaCl

- 0.05% NP-40
- 50 mM Tris-HCl, pH 7.4
- 2. G-50 buffer
 - 20 mM Tris, pH 7.5
 - 300 mM NaOAc
 - 2 mM EDTA, pH 8
 - 0.25% SDS
- 3. 1x TBE
 - 216 g Tris base
 - 110 g boric acid
 - 80 ml 0.5 M EDTA, pH 8
 - 1.5 L water
 - (Bring solution to pH 8.3)

I. Acknowledgments

We thank Dr. Joan Steitz's laboratory for sharing part of this protocol. I.M., J.A.B., and K.E.H. were supported in part by a WVU Foundation Fund (2V882). K.E.H. was supported in part by funding from the Ladies Auxiliary to the VFW of the United States (CK 003229). M.X. is supported by NIH grant R00 CA190886. J.A.S. is an Investigator of the Howard Hughes Medical Institute. This protocol was modified from previous work (Yu *et al.*, 1998). We do not have any conflict of interest or competing interests.

J. References

1. Boulon, S., Verheggen, C., Jady, B. E., Girard, C., Pescia, C., Paul, C., Ospina, J. K., Kiss, T., Matera, A. G., Bordonne, R. and Bertrand, E. (2004). [PHAX and CRM1 are required sequentially to transport U3 snoRNA to nucleoli.](#) *Mol Cell* 16(5): 777-787.
2. Collier, H. A. (2011). Cell biology. [The essence of quiescence.](#) *Science* 334(6059): 1074-1075.
3. Esteller, M. (2011). [Non-coding RNAs in human disease.](#) *Nat Rev Genet* 12(12): 861-874.
4. Ha, M. and Kim, V. N. (2014). [Regulation of microRNA biogenesis.](#) *Nat Rev Mol Cell Biol* 15(8): 509-524.
5. Kim, Y. K., Kim, B. and Kim, V. N. (2016). [Re-evaluation of the roles of DROSHA, Exportin 5, and DICER in microRNA biogenesis.](#) *Proc Natl Acad Sci U S A* 113(13): E1881-1889.
6. Kiss, T. (2004). [Biogenesis of small nuclear RNPs.](#) *J Cell Sci* 117(Pt 25): 5949-5951.
7. Luhrmann, R., Appel, B., Bringmann, P., Rinke, J., Reuter, R., Rothe, S. and Bald, R. (1982). [Isolation and characterization of rabbit anti-m3 2,2,7G antibodies.](#) *Nucleic Acids Res* 10(22): 7103-7113.

8. Martinez, I., Hayes, K. E., Barr, J. A., Harold, A. D., Xie, M., Bukhari, S. I. A., Vasudevan, S., Steitz, J. A. and DiMaio, D. (2017). [An Exportin-1-dependent microRNA biogenesis pathway during human cell quiescence.](#) *Proc Natl Acad Sci U S A* 114(25): E4961-E4970.
9. Ohno, M., Segref, A., Bachi, A., Wilm, M. and Mattaj, I. W. (2000). [PHAX, a mediator of U snRNA nuclear export whose activity is regulated by phosphorylation.](#) *Cell* 101(2): 187-198.
10. Palacios, I., Hetzer, M., Adam, S. A. and Mattaj, I. W. (1997). [Nuclear import of U snRNPs requires importin \$\beta\$.](#) *EMBO J* 16(22): 6783-6792.
11. Valcourt, J. R., Lemons, J. M., Haley, E. M., Kojima, M., Demuren, O. O. and Collier, H. A. (2012). [Staying alive: metabolic adaptations to quiescence.](#) *Cell Cycle* 11(9): 1680-1696.
12. Yedavalli, V. S. and Jeang, K. T. (2010). [Trimethylguanosine capping selectively promotes expression of Rev-dependent HIV-1 RNAs.](#) *Proc Natl Acad Sci U S A* 107(33): 14787-14792.
13. Yu, Y. T., Shu, M. D. and Steitz, J. A. (1998). [Modifications of U2 snRNA are required for snRNP assembly and pre-mRNA splicing.](#) *EMBO J* 17(19): 5783-5795.

Jamie Ann Barr

jamieabarr@gmail.com

Professional Summary

I aspire to work in a position that allows me to contribute to the translation of research discoveries to the public by utilizing my knowledge in technology transfer and commercialization, biology, and engineering principles. Experiences during my academic career and positions as a technology scout have shown me that I enjoy discussing scientific research topics, working with researchers to evaluate their technologies for commercial potential, and advising on the process of technology commercialization and protection.

Relevant Experience	<u>West Virginia University Health Sciences Innovation Center</u>	
	<u>Technology Scout</u>	February 2019-Present
	<u>Technology Scout Intern</u>	August 2017-January 2019
	<ul style="list-style-type: none"> ○ Serve as direct link between WVU researchers and Health Sciences Innovation Center to translate basic research to market ○ Attend research lab meetings to identify and discuss ideas for inventions and upcoming public disclosures ○ Assist researchers with new invention disclosure forms ○ Evaluate discoveries/ideas and advise on additional research activities to increase patentability and commercial viability ○ Perform prior art searches including peer reviewed literature and patents to evaluate novelty ○ Assist in the communication between researchers and outside legal counsel for the filing of patent applications ○ Develop portfolio with one-page highlights to market research inventions, resources, and investigators ○ Evaluate scientists' marketability reports from external market consulting firm ○ Familiarize with process of identifying licensing partners at a start-up company or existing company and educate researchers on partnership opportunities ○ Exposure to various agreements (e.g. MTA, NDA, Licensing) 	
	<u>Cancer Cell Biology Graduate Research Assistant at West Virginia University</u>	2011-2019
	Principal Investigator: Ivan Martinez, Ph.D.	Spring 2013-Spring 2019
	<ul style="list-style-type: none"> • Dissertation Thesis: Regulation of long non-coding RNAs by the human papillomavirus oncoproteins • Independently plan and conduct scientific experiments for multiple research projects • Plan projects to train multiple research rotation students • Perform weekly ordering for the lab based on lab budget • Maintain Radiation Safety compliance for the lab 	
	Principal Investigator: Yon Rojanasakul, Ph.D.	Spring 2012-Spring 2013
	<ul style="list-style-type: none"> • Researched the role of a cardiac glycoside as an anti-cancer therapeutic agent • Analyzed the toxicity of carbon nanotubes and their risk of inducing and promoting lung carcinogenesis 	
	Lectured a Cellular Methods Class to students in Ph.D. program	2014
	<u>Biomedical & Chemical Engineering Senior Research at West Virginia University</u>	2010-2011
	Principal Investigator: Robin S. Hissam, Ph.D.	
	<ul style="list-style-type: none"> • Performed research on elastin-based vesicles for the transport and delivery of drugs 	
	<u>Pharmaceutical Sciences Summer Research Experience at West Virginia University</u>	2010
	Principal Investigator: Peter M. Gannett, Ph.D.	

- Performed research using nanoparticles in order to improve detection and treatment of diseases

Education	<u>West Virginia University</u>	
	<u>School of Medicine</u>	August 2011-January 2019
	Doctor of Philosophy Candidate	
	Cancer Cell Biology	
	<u>West Virginia University</u>	
	<u>College of Engineering and Mineral Resources</u>	August 2008-May 2011
	Bachelor of Science in Chemical Engineering	
	<ul style="list-style-type: none"> • Graduated Magna Cum Laude • Certificate in Biomedical Engineering • Passed NCEES Fundamentals of Engineering Exam (Chemical) 	2011
Fellowships, Honors, and Scholarships	Graduate Student Travel Award to visit Broad Institute in Boston, MA	January 23-31, 2017
	Principal Investigator: John Rinn, Ph.D.	
	<ul style="list-style-type: none"> • Trained on RNA Immunoprecipitation technique 	
	Outstanding Graduate Student Award – Captain of Team Award	2016
	“Mad Scientists” - Relay for Life Team	
	1 st Place Van Liere Poster Presentation Award – Basic Science	
	Year 4	March 2016
	Year 3	February 2015
	WVNano/NanoSAFE Graduate Fellowship Program	August 2011-August 2013
	Dean’s List	Fall 2007 & 2009, Spring 2008, 2010, & 2011
	WVNano Summer Undergraduate Research Experience (SURE) program	April-July 2010
	Kincaid Memorial Scholarship	Fall 2009, Spring 2010

Skills	○ Completing and compiling documents for the disclosure of new technologies	
	○ Discussing and evaluating scientists' research ideas and guide inventors to advance their technologies	
	○ Identifying marketable attributes to researchers' inventions	
	○ Communication: oral presentation and written (assisted Principal Investigator on editing grant submissions from 2013-present) (writer, co-author, and editor on multiple scientific, technical writing journal submissions) (presented research at many student forums, conferences, and meetings)	
	○ Computer proficient: Microsoft Office Excel, PowerPoint, Word; Adobe Illustrator	
	○ Developing and optimizing detailed scientific protocols	
	○ RNA preparation and analysis for human transcriptome microarray (Affymetrix)	
	○ RNA-specific techniques: RNA & library preparation for high-throughput sequencing, RNA extraction, real-time reverse transcription PCR (q-RT-PCR), reverse transcription PCR (RT-PCR), PCR, RNA Immunoprecipitation, northern blotting, rapid amplification of cDNA ends (RACE), RNA fluorescence in situ hybridization (FISH) in cell lines and human tissue arrays	
	○ Experience with confocal microscopy	
	○ Genetic expression manipulation: CRISPR/Cas9, antisense oligonucleotides (ASOs), siRNA, shRNA	
	○ DNA extraction	
	○ Cellular Fractionation	

- Standard cell culture techniques (primary and immortalized mammalian cells)
- Mammalian cell retroviral and lentiviral transfection (Lipofectamine- and calcium phosphate-based) and infection
- Cloning and transformation of constructs in bacteria
- Restriction digestion & TOPO cloning
- *In vitro* cellular function assays (Transwell migration and invasion, flow cytometry for PI cell cycle and Annexin V/PI apoptosis, cell proliferation/viability)
- Protein-specific techniques: protein extraction, immunoprecipitation, western blotting

Publications

Manuscripts published:

Barr, J.A., Hayes, K.E., Brownmiller, T., Harold, A.D., Jagannathan, R., Lockman, P., Martinez, I. Long non-coding RNA FAM83H-AS1 is regulated by human papillomavirus 16 E6 independently of p53 in cervical cancer cells. *Sci Rep.* 2019 Mar 6;9(1):3662.

Hayes, K.E., **Barr, J.A.**, Xie, M., Steitz, J.A., Martinez, I. (2017) Immunoprecipitation of tri-methylated capped RNA. *Bio Protoc.* 2018 Feb 5;8(3).

Martinez, I., Hayes, K.E., **Barr, J.A.**, Harold, A., Xie, M., Bukhari, S.I.A., Vasudevan, S., Steitz, J.A., DiMaio, D. (2017) An Exportin-1-dependent microRNA Biogenesis Pathway During Human Cell Quiescence. *Proc Natl Acad Sci U S A.* 114(25):E4961-E4970.

Manuscripts in preparation for submission:

Khan, S.A., **Barr, J.A.**, Brownmiller, T., Martinez, I. Global Expression Analysis and Tissue-arrays Identify the Regulation of LAT1 and NAP1L3 Genes by Human Papillomavirus Oncogenes in Cervical-Cancer Cells. 2018 (Manuscript under review at *Journal of Virology*).

Hayes, K.E., **Barr, J.A.**, Harold, A., Wilusz, J.E., Martinez, I. Discovery of Novel Non-Coding Circular RNAs Generated by High-Risk Human Papillomaviruses. 2018.

Hayes, K.E., **Barr, J.A.**, Harold, A., Wilusz, J.E., Ziegelbauer, J.M., Martinez, I. HPV Oncogene E6 Regulates the Human Non-Coding Circular RNA circMYBL2 During Cervical Carcinogenesis. 2018.

Organizations	Association of University Technology Managers (AUTM)	December 2017-Present
	American Association for the Advancement of Science (AAAS)	August 2016-Present
	Cell Biology Training Program (CBPT)	Fall 2013-Spring 2019
	American Association for Cancer Research (AACR)	2012-Present
	Graduate Student Organization (GSO)	2012-2019
	○ Secretary	2014-2016
	Tau Beta Pi (TBP): The Engineering Honors Society	2010-Present
	First Year Student Mentorship Program	Fall 2013-Spring 2014
	WVNano/NanoSAFE Graduate Fellowship Program	August 2011-August 2013
	The American Institute of Chemical Engineers (AIChE)	2008-2011
	○ Secretary	Fall 2010-Spring 2011
	○ Assistant Treasurer	Fall 2009-Spring 2010
	Society of Women Engineers (SWE)	2008-2011
	Society of Biological Engineers (SBE)	2010-2011